



INTERNATIONAL CONGRESS ON
**PEER REVIEW
AND
BIOMEDICAL
PUBLICATION**

September 16-18, 2005
Chicago, Illinois, USA

JAMA
&
ARCHIVES
JOURNALS

BMJ
Publishing
Group

www.jama-peer.org

Table of Contents

Welcome	iv
Program	1
Poster Listings	5
Plenary Session Abstracts	
Friday, September 16	9
Saturday, September 17	15
Sunday, September 18	22
Poster Session Abstracts	
Saturday, September 17	31
Sunday, September 18	45
Sponsors	58
Exhibitors	60
Congress Organizers and Planners	61



FIFTH INTERNATIONAL CONGRESS ON PEER REVIEW AND BIOMEDICAL PUBLICATION

Welcome!

JAMA and the *BMJ* Publishing Group welcome you to Chicago and the Fifth International Congress on Peer Review and Biomedical Publication. This Congress follows the 4 highly successful Peer Review Congresses held in 1989 and 1993 in Chicago, in 1997 in Prague, and in 2001 in Barcelona. As before, we will have 3 days for presentations of new research into peer review and all aspects of scientific publication, including 45 plenary session presentations. Each plenary session research presentation will be followed by equal time for discussion and questions from the audience. In addition, there are 58 poster presentations scheduled for Saturday and Sunday. We hope you will take an active part in the program as we depend on your participation in the discussion sessions to make the Congress a success. Enjoy the Congress and enjoy Chicago!

Drummond Rennie, Congress Director
Annette Flanagin, Congress Coordinator

Fiona Godlee, European Director
Jane Smith, European Coordinator

Congress Advisory Board

Congress Director

Drummond Rennie
JAMA and Institute for
Health Policy Studies
University of California, San Francisco
San Francisco, California, United States

Congress Coordinator

Annette Flanagin
JAMA
Chicago, Illinois, United States

European Director

Fiona Godlee
BMJ
London, United Kingdom

European Coordinator

Jane Smith
BMJ
London, United Kingdom

John C. Bailar III
Washington, DC, United States

Lisa A. Bero
Institute for Health Policy Studies
University of California, San Francisco
San Francisco, California, United States

Floyd Bloom
The Scripps Research Institute
La Jolla, California, United States

Michael Callaham
University of California, San Francisco
San Francisco, California, United States

Frank Davidoff
Wethersfield, Connecticut,
United States

Catherine D. DeAngelis
JAMA
Chicago, Illinois, United States

Kay Dickersin
Brown University School of Medicine
Providence, Rhode Island, United States

Jeff Drazen
New England Journal of Medicine
Boston, Massachusetts, United States

Peter Gøtzsche
The Nordic Cochrane Center
Copenhagen, Denmark

Brian Haynes
McMaster University Faculty
of Health Sciences
Hamilton, Ontario, Canada

Richard Horton
The Lancet
London, United Kingdom

Donald Kennedy
Science and Stanford University
Stanford, California, United States

George D. Lundberg
Medscape
New York, New York, United States

Ana Marušić
Croatian Medical Journal
Zagreb, Croatia

David Moher
Children's Hospital of Eastern Ontario
Research Institute
Ottawa, Ontario, Canada

Daniel J. Ncayiyana
South African Medical Journal and
Durban Institute of Technology
Durban, South Africa

John Overbeke
Nederlands Tijdschrift voor Geneeskunde
Amsterdam, the Netherlands

Roy M. Pitkin
La Quinta, California, United States

Humberto B. Reyes
Revista Medica de Chile
Santiago, Chile

Povl Riis
Central Research Ethics of Denmark
University of Copenhagen
Hellerup, Denmark

Peush Sahni
National Medical Journal of India
New Delhi, India

Qian Shouchu
Chinese Medical Journal
Chinese Medical Association
Beijing, China

Richard Smith
UnitedHealth Europe Limited
London, United Kingdom

Harold C. Sox
Annals of Internal Medicine
Philadelphia, Pennsylvania,
United States

James K. Tumwine
African Health Sciences
Makerere University
Kampala, Uganda

Martin Van Der Weyden
Medical Journal of Australia
Pyrmont, New South Wales, Australia

Program

All plenary sessions will be held in the International Ballroom. Poster sessions will be held in the Gold Room. Breaks, Luncheons, and Exhibits will be held in the Imperial Ballroom.

FRIDAY, SEPTEMBER 16

8:00 AM - 9:00 AM

Registration, Continental Breakfast, and Exhibits

9:00 AM

Welcome and Introductions
Drummond Rennie (UNITED STATES)

9:10 AM - 9:40 AM

The Agony and the Ecstasy:
The History and Meaning of the Journal Impact Factor
Eugene Garfield (UNITED STATES)

9:40 AM - 10:20 AM

Authorship and Contributorship
Moderator: Drummond Rennie (UNITED STATES)

In the Eye of the Beholder: Contribution Disclosure
Practices and Inappropriate Authorship
Ana Marušić, Tamara Bates, Ante Anić, Vesna Ilakovac, and
Matko Marušić (CROATIA)

Declaration of Medical Writing Assistance in
International Peer-Reviewed Publications and Effect of
Pharmaceutical Sponsorship
Karen Woolley, Julie Ely, Mark Woolley, Felicity Lynch, Jane
McDonald, Leigh Findlay, and Yoonah Choi (AUSTRALIA,
JAPAN)

10:20 AM - 10:50 AM

Refreshment Break and Visit Exhibits

11:00 AM - 12:00 PM

Journal Guidelines and Policies
Moderator: Jane Smith (UNITED KINGDOM)

The Statistical and Methodological Content
of Journals' Instructions for Authors
Douglas G. Altman and David L. Schrager (UNITED STATES,
UNITED KINGDOM)

Questionnaire Availability From Published Studies in
3 Prominent Medical Journals
Lisa Schilling, Kristy Lundahl, and Robert Dellavalle
(UNITED STATES)

Conflict of Interest Disclosure Policies and Practices of
Peer-Reviewed Biomedical Journals
Richelle J. Cooper, Malkeet Gupta, Michael S. Wilkes, and
Jerome R. Hoffman (United States)

12:00 PM - 1:30 PM

Lunch and Visit Exhibits

1:30 PM - 2:30 PM

Peer Review Process
Moderator: Michael Callaham (UNITED STATES)

Editorial Changes to Manuscripts Published
in Major Biomedical Journals
Kirby P. Lee, Elizabeth A. Boyd, and Lisa A. Bero
(UNITED STATES)

Comparison of Author and Editor Suggested
Reviewers in Terms of Review Quality, Timeliness, and
Recommendation for Publication
Sara Schroter, Leanne Tite, Andrew Hutchings, and
Nick Black (UNITED KINGDOM)

Effect of Authors' Suggestions Concerning Reviewers
on Manuscript Acceptance
Lowell A. Goldsmith, Elizabeth Blalock, Heather Bobkova,
and Russell P. Hall (UNITED STATES)

2:30 PM - 3:30 PM

Peer Review Process
Moderator: John Overbeke (THE NETHERLANDS)

Assessment of Blind Peer Review on Abstract
Acceptance for Scientific Meetings
Joseph S. Ross, Cary P. Gross, Yuling Hong,
Augustus O. Grant, Stephen R. Daniels, Vladimir C.
Hachinski, Raymond J. Gibbons, Timothy J. Gardner,
and Harlan M. Krumholz (UNITED STATES)

Peering at Peer Review: Harnessing the Collective
Wisdom to Arrive at Funding Decisions About Grant
Applications
Nancy E. Mayo, James Brophy, Mark S. Goldberg,
Marina B. Klein, Sydney Miller, Robert Platt,
and Judith Ritchie (CANADA)

A Second Order of Peer Review:
Peer Review for Clinical Practitioners
Brian Haynes, Chris Cotoi, Leslie Walters,
Jennifer Holland, Nancy Wilczynski, Dawn Jedraszewski,
James McKinlay, and Ann McKibbin,
for the PLUS Project (CANADA)

3:30 PM - 4:00 PM

Refreshment Break and Visit Exhibits

4:00 PM - 5:00 PM

Scientific Misconduct

Moderator: Fiona Godlee (UNITED KINGDOM)

Retractions in the Research Literature: Misconduct or Mistakes?

Benjamin G. Druss, Sara Bressi, and Steven C. Marcus
(UNITED STATES)

Citation of Literature Flawed by Scientific Misconduct

A. Victoria Neale, Justin Northrup, Judith Abrams, and Rhonda Dailey (UNITED STATES)

For Which Cases of Suspected Misconduct Do Editors Seek Advice? An Observational Study of All Cases Submitted to the Committee on Publication Ethics (COPE)

Sabine Kleinert, Jeremy Theobald, Elizabeth Wager, and Fiona Godlee (UNITED KINGDOM)

6:00 PM - 7:30 PM

Welcome Reception

SATURDAY, SEPTEMBER 17

7:30 AM - 8:30 AM

Registration, Continental Breakfast, and Exhibits

8:30 AM - 9:00 AM

Journals—Agents for Change or Merely Mirrors on Society?

Richard Smith (UNITED KINGDOM)

9:00 AM - 10:20 AM

Publication Bias and Funding/Sponsorship

Moderator: Catherine D. DeAngelis (UNITED STATES)

Failures to Publish Pharmaceutical Clinical Trials

William Gardner and Charles W. Lidz (UNITED STATES)

Are Authors' Financial Ties With Pharmaceutical Companies Associated With Positive Results or Conclusions in Meta-analyses on Antihypertensive Medications?

Veronica Yank, Drummond Rennie, and Lisa A. Bero
(UNITED STATES)

Sponsorship, Bias, and Methodology: Cochrane Reviews Compared With Industry-Sponsored Meta-analyses of the Same Drugs

Anders W. Jørgensen and Peter C. Gøtzsche (DENMARK)

Is Everything in Health Care Cost-effective? Reported Cost-effectiveness Ratios in Published Studies

Chaim M. Bell, David R. Urbach, Joel G. Ray, Ahmed Bayoumi, Allison B. Rosen, Dan Greenberg, and Peter J. Neumann (CANADA, UNITED STATES, ISRAEL)

10:30 AM - 11:00 AM

Refreshment Break and Visit Exhibits

11:00 AM - 12:00 PM

Publication Bias and Journal Factors

Moderator: Peush Sahni (INDIA)

Characteristics of Accepted and Rejected Manuscripts at Major Biomedical Journals: Predictors of Publication

Kirby P. Lee, Elizabeth A. Boyd, Peter Bacchetti, and Lisa A. Bero (UNITED STATES)

Rethinking Publication Bias: Developing a Schema for Classifying Editorial Discussion

Kay Dickersin and Catherine Mansell (UNITED STATES)

Is Publication Bias Associated With Journal Impact Factor?

Yuan-I Min, Aynur Unalp-Arida, Roberta Scherer, and Kay Dickersin (UNITED STATES)

12:00 - 1:30 PM

Lunch and Visit Exhibits

1:30 PM - 2:30 PM

Effect of Indexing, Open Access, and Journal "Phenomena" on Submissions, Citations, and Impact Factors

Moderator: Annette Flanagan (UNITED STATES)

Impact of SciELO and MEDLINE Indexing on the Submission of Articles to a Non-English-Language Journal

Danilo Blank, Claudia Buchweitz, and Renato S. Procianoy
(BRAZIL)

Effect of Open Access on Citation Rates for a Small Biomedical Journal

Dev Kumar R. Sahu, Nithya J. Gogtay, and Sandeep B. Bavdekar (INDIA)

More Than a Decade in the Life of the Impact Factor

Mabel Chew, Martin Van Der Weyden, and Elmer V. Villanueva (AUSTRALIA)

2:30 PM - 3:30 PM

Poster Sessions, Refreshments, and Visit Exhibits

(See Poster listings on pages 5-6 for details)

3:30 PM - 4:50 PM

Quality and Influence of Citations and Dissemination of Scientific Information to the Public

Moderator: Martin Van der Weyden (AUSTRALIA)

Use and Persistence of Internet Citations in Scientific Publications: An Automated 5-Year Case Study of Dermatology Journals

Kathryn R. Johnson, Jonathan D. Wren, Lauren F. Heilig, Eric J. Hester, David M. Crockett, Lisa M. Schilling, Jennifer M. Myers, Shayla Orton Francis, and Robert P. Dellavalle (UNITED STATES)

Are Reviewers Influenced by Citations of Their Own Work? Evidence From the *International Journal of Epidemiology*

Matthias Egger, Lesley Wood, Erik von Elm, Anthony Wood, Yoav Ben Shlomo, and Margaret May (UNITED KINGDOM, SWITZERLAND)

How the News Media Report on Research Presented at Scientific Meetings: More Caution Needed

Steven Woloshin and Lisa M. Schwartz (UNITED STATES)

Reading Between the Ads: Assessing the Quality of Health Articles in Top Magazines

Brad Hussey, Diane Miller, and Alejandro R. Jadad (CANADA)

World Association of Medical Editors Meeting

Saturday, September 17, 2005

5:15 PM to 7:15 PM

The officers of the World Association of Medical Editors (WAME) invite all registrants to the next WAME meeting. The meeting will include a discussion (with audience participation) of some common, real-life ethical problems faced by editors, by panelists expert in editorial and publication ethics issues, as well as a business meeting. WAME is a voluntary association of more than 1200 editors from 85 countries who seek to foster cooperation among editors of peer-reviewed medical journals worldwide, and to provide resources and guidelines to help editors raise the standards of scientific publication. For more information about WAME, see <http://www.wame.org>.

SUNDAY, SEPTEMBER 18

7:30 AM - 8:30 AM

Registration, Continental Breakfast, and Exhibits

8:30 AM - 9:00 AM

More Peer Review Congresses? Why?

Drummond Rennie (UNITED STATES)

9:00 AM - 10:00 AM

Reporting of Studies: Abstracts and Publication After Meeting Presentations

Moderator: David Moher (CANADA)

Trials Reported in Abstracts: The Need for a Mini-CONSORT

Sally Hopewell and Mike Clarke (UNITED KINGDOM)

Are Relative Risks and Odds Ratios in Abstracts Believable?

Peter C. Gøtzsche (DENMARK)

Do Clinical Trials Get Published After Presentation at Biomedical Meetings?

A Systematic Review of Follow-up Studies

Erik von Elm and Roberta Scherer (SWITZERLAND, UNITED STATES)

10:00 AM - 10:30 AM

Refreshment Break and Visit Exhibits

10:30 AM - 11:50 AM

Improving the Quality of Reporting of Trials and Other Studies

Moderator: Ana Marušić (CROATIA)

Design, Analysis, and Presentation of Crossover Trials

Edward J. Mills, An-wen Chan, Ping Wu, Gordon H. Guyatt, and Douglas G. Altman (CANADA, UNITED KINGDOM)

Quality of Trials in Operative Surgery:

Where Is the Comic Opera?

Catherine Jane Walter, Jo Dumville, Catherine Hewitt, David Torgerson, Philip Drew, and John R. T. Monson (UNITED KINGDOM)

Does the CONSORT Checklist Improve the Quality of Reports of Randomized Controlled Trials?

A Systematic Review

Amy C. Plint, David Moher, Kenneth Schulz, Douglas G. Altman, and Andra Morrison (CANADA, UNITED STATES, UNITED KINGDOM)

Does the STARD Statement Improve the Quality of Reporting of Diagnostic Accuracy Studies?
Nynke Smidt, Anne W. S. Rutjes,
Daniëlle A. W. M. van der Windt, Raymond W. J. G. Ostelo,
Johannes B. Reitsma, Patrick M. Bossuyt, Lex M. Bouter,
and Henrica C. W. de Vet (THE NETHERLANDS)

12:00 PM - 1:30 PM

Lunch and Visit Exhibits

1:30 PM - 2:50 PM

Quality of Reporting of Trials: Protocols, Manuscripts, Published Articles, and Postpublication
Moderator: Kay Dickersin (UNITED STATES)

Comparison of Submitted and Published Reports of Randomized Trials

Douglas G. Altman, John P. A. Ioannidis, David Moher,
Jill Mollison, David L. Schriger, and Sara Schroter
(UNITED KINGDOM, GREECE, CANADA, UNITED STATES)

Reporting of Study Outcomes in Comparative Drug Trials: Evidence From Proposals Submitted to a Research Ethics Committee and Corresponding Full Publications

Erik von Elm, Karin Huwiler, Mark Witschi,
Alexandra Röllin, Charles Senessie, Nicola Low,
and Matthias Egger (SWITZERLAND, UNITED KINGDOM)

A Comparison of the Published Version of Randomized Controlled Trials in a Specialist Clinical Journal With the Original Trial Protocols

Lucy Chappell, Zarko Alfirevich, Patrick Chien, Susan Jarvis,
and Jim G. Thornton (UNITED KINGDOM)

Trial Bank Publishing of Randomized Trials: Preliminary Results
Ida Sim and Ben Olasov (UNITED STATES)

3:00 PM - 4:00 PM

Poster Sessions and Refreshments
(See Poster listing on pages 6-7 for details)

4:00 PM - 5:00 PM

Hot Topics: Open Access and Trial Registries
Moderator: Drummond Rennie (UNITED STATES)

Authors' Access to Financial Support at the Time of Paper Acceptance: A Survey of Biomedical Journals
Sara Schroter, Leanne Tite, and Ahmed Kassem
(UNITED KINGDOM)

Would You Drop Your Membership? Professional Organization Members' Reaction to Open Access
Harold C. Sox and Wayne Bylsma (UNITED STATES)

Assessing the Quality of Information Recorded on Trial Registries

Munira Nurbhai, Pasquale Lorenzo Moja, Jeremy Grimshaw,
Alessandro Liberati, An-Wen Chan, Kay Dickersin,
Karmela Krleza-Jeric, David Moher, Ivan Moschetti,
Drummond Rennie, Ida Sim, and Jimmy Volmink
(CANADA, ITALY, UNITED STATES, SOUTH AFRICA)

5:00 PM

Summary and Adjournment
Drummond Rennie

Poster Listings

All posters will be presented in the Gold Room

SATURDAY, SEPTEMBER 17

Authorship and Contributorship

Quantification of Author Contributions and Eligibility for Authorship: A Randomized Trial
Ana Ivaniš, Darko Hren, Aleksandra Mišak, Dario Sambunjak, Matko Marušić, and Ana Marušić (CROATIA)

Conflict of Interest

Conflict of Interest Policies at Scientific Journals: A Cross-Disciplinary Comparison
Jessica S. Ancker and Annette Flanagan (UNITED STATES)

Conflicts of Interest Disclosed by Authors of Manuscripts Submitted to a General Medical Journal
Christine Laine, Mary Beth Schaeffer, and Catharine Stack (UNITED STATES)

Ethical Concerns

Imaging and Non-imaging Journal Policies Regarding Institutional Review Board Approval and Informed Consent Declarations by Authors
Andrew Y. Choi, Douglas S. Katz, and Anthony V. Proto (UNITED STATES)

Reporting of Ethical Committee Approval and Patient Consent by Study Design in 5 General Medical Journals
Sara Schroter, Ros Plowman, Andrew Hutchings, and Adrian Gonzalez (UNITED KINGDOM)

Instructions for Authors

The Impact of Editorial Guidelines on the Classification of Race/Ethnicity in the *BMJ*
George T. H. Ellison, Mikey Rosato, and Simon Outram (UNITED KINGDOM)

An Analysis of the Content of Medical Journals' Instructions for Authors
David L. Schrager, Sanjay Arora, Varda A. Schrager, and Douglas G. Altman (UNITED STATES, UNITED KINGDOM)

Detailed Instructions to Authors Are Required in Chinese Medical Journals
Qian Shou-chu (CHINA)

Open Access

The Impact of Author Page Charges on Published Research in Infectious Diseases
Surabhi Liyanage and C. Raina MacIntyre (AUSTRALIA)

Peer Review Process

Do You Need to Be an Editor to Accept or Reject Research Papers Sent to the *BMJ*?
Gary Bryan, John Fletcher, and Rajendra Kale (UNITED KINGDOM)

Qualitative Profile of Journal Peer Reviewers and Predictors of Peer Reviewer Quality
Michael Callahan and John Tercier (UNITED STATES)

A Randomized Trial on the Effect of Statistical Reviewing and Checklist on Manuscript Quality
Erik Cobo, Albert Selva, Josep Maria Ribera, Francesc Cardellach, Ruth Dominguez, Agustín Urrutia, Vicens Fonollosa, Mercedes Belmonte, Celestino Rey-Joly, and Miquel Vilardell (SPAIN)

Blinded vs Unblinded Peer Review in a Non-English-Language Journal: A Randomized Controlled Trial
Torben V. Schroeder, and Ole H. Nielsen, for the Editorial Team of *Ugeskrift for Læger (Journal of Danish Medical Association)* (DENMARK)

Consistency Between Reviewers and Editors About Which Papers Should Be Published
James R. Scott, Sheryl Martin, and Leon Burmeister (UNITED STATES)

Are Reviewers Suggested by Authors as Good as Those Chosen by Editors? Results of a Rater-Blinded, Retrospective Study
Elizabeth Wager, Emma C. Parkin, and Pritpal S. Tamber (UNITED KINGDOM)

Publication Bias

How Does Prior Publication Affect Full Publication of Completed Clinical Trials?
Yuan-I Min, Aynur Unalp-Arida, Roberta Scherer, and Kay Dickersin (UNITED STATES)

Quality of Reporting Trials and Other Studies

What's in a NAME: Clinical Trial Acronyms and Research Impact
Matthew B. Stanbrook and Donald A. Redelmeier (CANADA)

Quality of Placebo-Controlled Trials of Alternative and Conventional Medicine: Matched-Pair Study
Aijing Shang, Karin Huwiler, Linda Nartey, Peter Jüni, and Matthias Egger (SWITZERLAND, UNITED KINGDOM)

Improving the Quality of Reporting of Randomized Clinical Trials Evaluating Herbal Interventions: An Extension of the CONSORT Statement
Joel J. Gagnier, Heather Boon, Paula Rochon, David Moher, Joanne Barnes, and Claire Bombardier, for the CONSORT Group (CANADA)

An Analysis of the Quality of Reporting of General Surgical Randomized Controlled Trials Published in General Health Care and Surgical Journals
Martin Wiener, Sabapathy Prakash Balasubramanian, Zeiad Kaid, Ravindranath Tiruvoipati, Diana Elbourne, and Malcolm Walter Reed (UNITED KINGDOM)

Quality of Journal Articles

Titles of Articles in Peer-Reviewed Journals Lack Essential Information: A Structured Review of Contributions to 4 Leading Medical Journals, 1995 and 2001
Paul Z. Siegel, Stephen B. Thacker, Richard A. Goodman, and Cathleen Gillespie (UNITED STATES)

References and Citations

Editors' Impact on Improving the Accuracy of References: Randomized Comparison of Standard Practice, Brief Reminder, or Educational Intervention
Kristina Fišter, Ana Marušić, Andrew Hutchings, Josipa Kern, and Matko Marušić (CROATIA)

Web Citations: Going, Going, Still There
Gunther Eysenbach, M. J. Suhonos, and Jean-Sebastian Dumais (CANADA)

Rhetoric and Interpretation

The Rhetoric of Efficacy in Donepezil Trials
John R. Gilstad and Thomas E. Finucane (UNITED STATES)

Do the Conclusions Look as Good as They Seem? A Review of Quality Improvement Intervention Studies
Linda Li, Lorenzo Moja, Alberto Romero, and Jeremy Grimshaw (CANADA, ITALY, SPAIN)

Web and e-Publishing

The Effect of Using e-Mail Push Technology on Readership of Articles in a General Medical Journal
George D. Lundberg and Kaytie Brown (UNITED STATES)

Prepublication Release of Journal Articles: Impact of the Electronic Medium on the Research Message
Matthew B. Stanbrook and Donald A. Redelmeier (CANADA)

SUNDAY, SEPTEMBER 18

Authorship and Contributorship

How Intuitive are ICMJE Criteria for Authorship? Perceptions of Deserved Authorship Among Medical Students and Physicians
Darko Hren, Dario Sambunjak, Ana Ivaniš, Matko Marušić, and Ana Marušić (CROATIA)

Conflict of Interest

Conflict of Interest Management at the International CPR Consensus Conference
John E. Billi, Brian Eigel, David Zideman, and Vinay Nadkarni (UNITED STATES)

Duplicate Publication

The Extent of Duplicate Publications in a Cohort of Meta-analyses
Veronica Yank, Drummond Rennie, and Lisa A. Bero (UNITED STATES)

Ethical Concerns

Assessment of Equipoise Using a Cohort of Randomized Controlled Trials
Yuan-I Min, Aynur Unalp-Arida, Roberta Scherer, and Kay Dickersin (UNITED STATES)

Has Reporting on Informed Consent, Ethical Approval, Competing Interest, and Financial Support Been Improved in Randomized Controlled Trial Articles in 3 Chinese Medical Journals?
Qian Shou-chu, Lv Xiao-dong, and Liu Bin (CHINA)

Impact Factor

Text vs Context: The Influence of the Journal on Article Impact
Matthew B. Stanbrook and Donald A. Redelmeier (CANADA)

Open Access

Open Access Deposition Policies: What Does It Mean for Medical Journals?

Serena J. Cubie, Gavriel A. Hollander, Sarah C. Price, and Richard A. Watt (UNITED KINGDOM)

Open Source Tools for Open Access Publishing

M. J. Suhonos and Gunther Eysenbach (CANADA)

Peer Review Process

Screening Parameters for Reviewer Selection

Michael Callaham and John Tercier (UNITED STATES)

Factors Affecting the Time From Manuscript Submission to Manuscript Acceptance

Julie Ely, Mark Woolley, Felicity Lynch, Jane McDonald, Leigh Findlay, Yoonah Choi, and Karen Woolley (AUSTRALIA, JAPAN)

Does Consumer Refereeing Improve the Quality of Systematic Reviews of Health Care Interventions?

The Perspectives of Editors and Authors

Gill Gyte, Carol Grant-Pearce, Sonja Henderson, Dell Horey, Sandy Oliver, and Carol Sakala (UNITED KINGDOM, AUSTRALIA, UNITED STATES)

Quality Measurement of Reviewers' Reports by a Simple Instrument

Annemieke P. Landkroon, Hans Veeken, Peter Hart, and A. John P. M. Overbeke (THE NETHERLANDS)

Behaviors of Authors and Peer Reviewers Following Change From Closed to Open System

L. Michael Posey (UNITED STATES)

Continuing Medical Education Credit as an Incentive for Participation in Peer Review

Mary Beth Schaeffer, Christine Laine, and Catharine Stack (UNITED STATES)

Why Do Peer Reviewers Decline to Review? A Survey

Leanne Tite and Sara Schroter (UNITED KINGDOM)

Characteristics of Reviews for a Series of Open Access Medical Journals: Results of a Retrospective Study

Elizabeth Wager, Emma C. Parkin, and Pritpal S. Tamber (UNITED KINGDOM)

Publication Bias

Constraints on Academic Freedom in Industry-Initiated Clinical Trials

Peter C. Gøtzsche, Asbjørn Hróbjartsson, Helle Krogh Johansen, Mette Haahr, Douglas G. Altman, and An-Wen Chan (DENMARK, UNITED KINGDOM, CANADA)

Quality of Journal Articles

From Submission to Publication: A Study of the Tables and Figures in a Cohort of RCTs Submitted to the *BMJ*

David L. Schriger, Reshmi Sinha, Pamela Liu, Douglas G. Altman, and Sara Schroter (UNITED STATES, UNITED KINGDOM)

Quality of Reporting Trials and Other Studies

Reporting Methods of Adverse Events in Randomized Controlled Trials

Curtis Sather and Jim Nuovo (UNITED STATES)

Grading the Evidence of Published Papers for the Benefit of Clinicians

James R. Scott, Rebecca Rinehart, and Catherine Y. Spong (UNITED STATES)

Statistics

Independent Re-analyses of Identical Data Sets: Implications for Peer Review

Penelope J. Greene (UNITED STATES)

Reliability of 3 Types of Research Methodologies in Obstetrics

Kerry M. McMahon, D. Yvette LaCoursiere, and James R. Scott (UNITED STATES)

Ratio Measures in Leading Medical Journals: Where Are the Underlying Absolute Risks?

Lisa M. Schwartz, Steven Woloshin, Evan L. Dvorin, and H. Gilbert Welch (UNITED STATES)

Web and e-Publishing

Inviting Conversation: Engaging Diverse Individuals and Groups in an Interactive Online Forum on Published Research

Laura A. McLellan, Robin S. Gotler, and Kurt C. Stange (UNITED STATES)

The Use of the World Wide Web by Medical Journals

David L. Schriger, Sripha Ouk, and Douglas G. Altman (UNITED STATES, UNITED KINGDOM)

Notes

Plenary Session Abstracts

All plenary sessions will be held in the International Ballroom.

FRIDAY, SEPTEMBER 16

Authorship and Contributorship

In the Eye of the Beholder: Contribution Disclosure Practices and Inappropriate Authorship

Ana Marušić,^{1,2} Tamara Bates,² Ante Anić,³ Vesna Ilakovac,⁴ and Matko Marušić^{1,2}

Objective To determine the effects of the structure of contribution disclosure forms on the number of authors not meeting criteria of the International Committee of Medical Journal Editors (ICMJE) and to analyze authors' contributions for the same article declared first by the corresponding author and then by individual authors.

Design In a single-blind randomized controlled trial, 1462 authors of 332 manuscripts submitted to the *Croatian Medical Journal* were sent 3 different contribution disclosure forms: open-ended; categorical, with 11 possible contribution choices; and instructional, instructing how many contributions are needed to satisfy individual ICMJE criteria. Main outcome measure was the number of authors not satisfying ICMJE criteria (honorary authors). In a separate study, corresponding authors of 201 submitted articles, representing 919 authors, received contribution disclosure forms with 11 possible contribution choices to declare contributions of all authors. The same form was then sent to each individual author, including the corresponding author.

Results In the randomized trial, the group answering the instructional form had significantly fewer authors whose reported contributions did not satisfy ICMJE criteria (18.7%) than did groups answering categorical (62.8%) or open-ended (54.7%) forms ($\chi^2_2 = 210.8$, $P < .001$). All authors answering the open-ended form, regardless of their compliance with authorship criteria, reported significantly fewer contributions (median, 3; 95% confidence interval [CI], 3-3) than did authors responding to either the categorical (median, 4; 95% CI, 4-4; z score = -7.19; $P < .001$) or instructional forms (median, 4; 95% CI, 4-5; z score = -13.98; $P < .001$). Honorary authors answering instructional forms reported more contributions (median, 3; 95% CI, 3-3) than those answering either the categorical (median, 2; 95% CI, 2-3; z score = 2.76; $P = .006$) or open-ended forms (median, 2; 95% CI, 2-2; z score = 3.38; $P < .001$). Most honorary authors (39.9%) lacked only the third ICMJE criterion (final approval of manuscript). In the second study, 201 (28.0%) of 718 noncorresponding authors met all 3 ICMJE criteria according to corresponding author's statement compared with 287 (40.0%) individual author's declarations (exact McNemar test, $S1 = 48.7$; $P < .001$). Disclosure forms filled out twice by corresponding authors for the same manuscript disagreed in 140 cases (69.3%).

Conclusions The structure of contribution disclosure forms significantly influences the number of contributions reported by authors and their compliance with ICMJE authorship criteria. Discrepancy between what corresponding authors declare on 2 separate occasions about their own contributions to the same manuscript indicate that recall bias may be one of the confounding factors of responsible authorship practices. Journal editors should

be aware of the cognitive aspects of survey methodology when they construct self-reports about behavior, such as contribution disclosure forms.

¹Croatian Medical Journal, Zagreb, Croatia; ²Zagreb University School of Medicine, Salata 3, HR-10000 Zagreb, Croatia, e-mail: marusica@mef.hr; ³Holy Ghost General Hospital, Zagreb, Croatia; ⁴Josip Juraj Strossmayer University, School of Medicine, Osijek, Croatia

Declaration of Medical Writing Assistance in International, Peer-Reviewed Publications and Effect of Pharmaceutical Sponsorship

Karen Woolley,^{1,2} Julie Ely,² Mark Woolley,² Felicity Lynch,² Jane McDonald,³ Leigh Findlay,² and Yoonah Choi²

Objective Medical writing assistance may improve manuscript quality and timeliness. Good Publication Practice Guidelines for pharmaceutical companies encourage authors to acknowledge medical writing assistance. The objectives of this study were to determine the proportion of articles from international, high-ranking, peer-reviewed journals that declared medical writing assistance and to explore the association between pharmaceutical sponsorship and medical writing assistance in terms of time to manuscript acceptance.

Design The acknowledgment sections of 1000 original research articles were reviewed. The sample comprised 100 consecutive articles published up to January 2005 from each of 10 high-ranking (impact factor-based), international, peer-reviewed medical journals from different therapeutic areas. The proportion of articles declaring pharmaceutical sponsorship and medical writing assistance, and the time interval between manuscript submission and acceptance were calculated. Analysis of variance was used to explore associations between sponsorship, writing assistance, and manuscript acceptance time.

Results Medical writing assistance was declared in only 6% of publications ($n = 60$). In the pharmaceutical-sponsored studies subset ($n = 102$ articles), assistance was declared in 10 articles (10%). Disclosure of medical writing assistance was associated with reduced time to acceptance (declared assistance: geometric mean, 83.6 days; vs no declaration: geometric mean, 132.2 days; relative difference, 0.63; 95% confidence interval, 0.40-1.01; $P = .053$).

Conclusions Based on this 1000-article sample, the reported use of medical writing assistance appears low (6%). If used, medical writing assistance should be declared. For pharmaceutical-sponsored studies, the time to publication may be faster when medical writers are used.

¹University of Queensland, 8 Shipyard Circuit, Noosaville, Queensland 4566, Australia, e-mail: kw@proscribe.com.au; ²ProScribe Medical Communications, Queensland, Australia; ³ProScribe Medical Communications, Tokyo, Japan

Journal Guidelines and Policies

The Statistical and Methodological Content of Journals' Instructions for Authors

Douglas G. Altman¹ and David L. Schrager^{1,2}

Objective To characterize the methodological and statistical advice provided by the instructions for authors of major medical journals.

Design Using citation impact factors for 2001, we identified the top 5 journals from each of 33 medical specialties and the top 15 from general and internal medicine that publish original clinical research. The final sample of 166 journals was obtained after examining 232 journals (some journals represent more than 1 specialty). We obtained the online instructions for authors for each journal between January and May 2003 and identified all material relating to 15 methodological and statistical topics. Assessments were performed by reading the instructions for authors and validated by text searches for relevant words.

Results Fewer than half the journals provided any information on statistical methods (TABLE 1). General journals were more likely to refer to reporting guidelines but less likely to include advice about statistical issues. Few journals (13%) commented on the use of tables and figures. Those instructions that referenced methodology papers cited from 1 to 49 (median, 1). There were many contradictions among instructions (eg, "Report actual *P* values, rather than ranges or limits" vs "Statistical probability (*p*) should be reported ... at only one of the following levels

$p < 0.05, 0.01, 0.005, \text{ and } 0.001$ " and "tables usually convey more precise numerical information; graphs should be reserved for highlighting changes over time or between treatments" vs "For presentation of data, figures are preferred to tables"). Other instructions were opaque (eg, "In general, statistical treatment of multiple experiments is required and is preferred to representative experiments; we prefer to see the SD unless sets of data from experiments or groups are pooled, in which case the SEM may be used").

Conclusions Journal instructions for authors provide little guidance regarding methodological and statistical issues and the advice provided may be unhelpful and at times contradictory.

¹Cancer Research UK/NHS Centre for Statistics in Medicine, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, UK, e-mail: doug.altman@cancer.org.uk; ²University of California Los Angeles Emergency Medicine Center, University of California Los Angeles School of Medicine, Los Angeles, CA, USA

Questionnaire Availability From Published Studies in 3 Prominent Medical Journals

Lisa Schilling, Kristy Lundahl, and Robert Dellavalle

Objective To describe the availability of questionnaires used for research published in 3 high-circulation medical journals.

Design A MEDLINE search with OVID using the terms "questionnaire" or "survey" identified 368 putative studies published between January 2000 and May 2003 in *JAMA*, *New England Journal of Medicine*, or *The Lancet*. Studies using nonnovel questionnaire instruments (eg, CAGE, Behavioral Risk Factor Surveillance Survey) were excluded. For inclusion, 2 investigators independently judged that the results of a questionnaire constituted the main outcome of the study. For qualifying studies with duplicate corresponding authors, 1 study was randomly selected for inclusion. Eighty-seven qualifying studies remained. Five publications contained reproductions of the administered questionnaire. In June 2004, corresponding authors of the remaining 82 articles were asked via written correspondence to provide a copy of the questionnaire used in their study. Nonresponders were mailed 2 additional requests at 2-week intervals.

Results Forty-four (54%; 95% confidence interval [CI], 43%-64%) of the requested questionnaires were provided; 38 (46%; 95% CI, 36%-57%) of the requested questionnaires were not provided. Of those not provided, no reason was provided for 26. The following other responses were provided: no access to the questionnaire ($n = 3$), questionnaire only given to collaborators ($n = 2$), referred to questionnaire description in the article ($n = 2$), change in job status ($n = 1$), no structured questionnaire used ($n = 1$), returned a questionnaire unrelated to the paper ($n = 1$), returned partial questionnaire ($n = 1$), and correct e-mail or postal address unavailable ($n = 1$).

Conclusions Our study shows that 6% of novel questionnaires were published, and only 54% of unpublished questionnaires could be obtained via the corresponding author. Reforms, such as mandatory publication or deposition in an open access

Table 1. Journal Information for Authors on Statistical and Methodological Issues

Instructions Address	No. (Range) for All Journals (n = 166)	Proportion of Instructions With This Item, %		Difference (95% CI)
		General (n = 15)	Specialty (n = 151)	
ICMJE Uniform Requirements	41 (33-49)	87	36	50 (24 to 62)
CONSORT	22 (16-29)	53	19	35 (11 to 57)
Other guidelines (QUOROM, STARD, etc)	4 (2-8)	33	1	32 (14 to 57)
"Provide enough detail for reproducibility"	11 (7-17)	0	12	-12 (-18 to 9)
Anything about statistical methods	39 (32-47)	27	40	-14 (-31 to 13)
<i>P</i> values	20 (15-27)	20	21	-1 (-15 to 25)
Confidence intervals	14 (9-21)	7	15	-9 (-17 to 15)
Multiple testing	5 (3-10)	0	6	-6 (-11 to 15)
Modeling	4 (2-8)	0	5	-5 (-9 to 16)
Power	10 (6-16)	0	11	-11 (-17 to 10)
Sensitivity analysis	4 (2-8)	0	5	-5 (-18 to 10)
Bayesian methods	1 (0-4)	7	1	6 (0 to 29)
Tables and figures	13 (8-19)	7	13	-7 (-20 to 7)
Statistical consultants	6 (3-11)	7	6	1 (-7 to 24)
Reference methods/statistics papers	10 (6-16)	33	8	25 (6 to 51)

Abbreviations: CI, confidence interval; ICMJE, International Committee of Medical Journal Editors; CONSORT, Consolidated Standards of Reporting Trials; QUOROM, Quality of Reporting of Meta-analyses; STARD, Standards for Reporting Diagnostic Accuracy.

archive, are needed to guarantee access to the fundamentally indispensable component of questionnaire research, the questionnaire.

Departments of Medicine and Dermatology, University of Colorado at Denver and Health Sciences Center, 2400 E 9th Ave, B-180, Denver, CO 80262, USA, e-mail: lisa.schilling@uchsc.edu

Conflict of Interest Disclosure Policies and Practices of Peer-Reviewed Biomedical Journals

Richelle J. Cooper,¹ Malkeet Gupta,¹ Michael S. Wilkes,² and Jerome R. Hoffman¹

Objective We undertook this investigation to characterize the conflict of interest (COI) policies of biomedical journals with respect to authors, peer-reviewers, and editors and to ascertain what information about COI disclosures is publicly available.

Design We performed a cross-sectional survey of a convenience sample of 135 peer-reviewed biomedical journal editors. We included a broad range of North American and European, general and specialty medical journals that publish original, clinical research topics, based on impact factor, and the recommendations of experts in the field. We reviewed each journal's Web page to identify the editors, and each editor in our sample represented a single journal only. We developed and pilot tested a 3-part Web-based survey prior to data collection. The survey included questions about the presence of specific policies for authors, peer-reviewers, and editors, any specific restrictions on authors, peer-reviewers, and editors based on COI, and the public availability of these disclosures. We contacted the journal editor's with a minimum of 3 requests, to improve response rate, and provided a written version of the survey for those unable to access the Web site.

Results The response rate for the survey was 91/135 (67%). Eighty-five (93%) journals have an author COI policy. Ten (11%) journals restrict author submissions based on COI (eg, drug company authors' papers on their products are not accepted). While 77% report collecting COI information on all author submissions, only 57% publish all author disclosures. Of journal respondents, 42/91 (46%) and 36/91 (40%) have a specific policy on peer-reviewer and editorial COI, respectively; 25% and 31% of journals require recusal of peer-reviewers and editors, respectively, if they report a COI; and 3% publish peer-reviewer COI disclosures and 12% publish editor COI disclosures, but 11% and 24% respectively, report the information is available on request.

Conclusions While most journals in our sample report having an author COI policy, the disclosures are not collected or published universally. Journals less frequently reported COI policies for peer reviewers and editors, and even less commonly publish those disclosures. Specific policies to restrict authorship, or peer-review and editing, were uncommonly reported in our sample.

¹UCLA Emergency Medicine Center, UCLA School of Medicine, 924 Westwood Blvd, #300, Los Angeles, CA, USA, e-mail: richelle@ucla.edu;

²UC Davis School of Medicine, CA, USA

Peer Review Process

Editorial Changes to Manuscripts Published in Major Biomedical Journals

Kirby P. Lee, Elizabeth A. Boyd, and Lisa A. Bero

Objective To assess the added value of peer review by identifying and characterizing editorial changes made to initial submissions of manuscripts that were accepted and published.

Design A prospective cohort of original research articles (n = 1107) submitted to the *Annals of Internal Medicine*, *BMJ*, and *Lancet* between January 2003 and April 2003. Experimental and observational studies, systematic reviews, and qualitative, ethnographic, or nonhuman studies were included. Single case reports were excluded. The original submitted manuscript was compared with the final publication. Changes were documented separately for abstract and text and qualitatively categorized. Minor changes were classified as additions or deletions of words that improved the clarity, readability, or accuracy of reporting. Major changes were classified as additions or deletions of data, new or revised statistical analyses, or statements affecting the interpretation of results or conclusions. We also documented changes in authorship and disclosure of any funding source (including its role) or potential conflicts of interest. Proportions and frequencies were calculated using the manuscript as the unit of analysis.

Results Of 1107 submitted manuscripts, 68 (6.1%) were accepted for publication. Changes from submission to publication are listed in TABLE 2. Extreme variability in the editorial process was observed ranging from near-verbatim publication of submitted manuscripts to extensive additions and deletions of data or text and revised statistical analyses with new tables and figures. Most changes were minor consisting of simple word modifications, statements, and clarifications. Major changes often included toning down the original Conclusions, emphasizing study limitations, or revising statistical analyses. Disclosure of potential conflicts of interest, any funding source, and its role in the design, conduct, and publication process improved in the final publication.

Table 2. Editorial Changes to Submitted Manuscripts

Editorial Changes	No./Total (%)	
	Abstract (n = 61)*	Text (n = 68)
Major	26/61 (43)	45/68 (66)
Minor	55/61 (90)	67/68 (99)
None	5/61 (8)	1/68 (1)
Added author		4/68 (6)
Deleted author		3/68 (4)
Change in order of authors		2/68 (3)
	Submission (n = 68)	Publication (n = 68)
Disclosure of any funding source	53/68 (78)	66/68 (97)
Disclosure of conflicts of interest	26/68 (38)	68/68 (100)
Described role of funding source†	10/54 (19)	41/54 (76)

*Not all manuscripts contained an abstract.

†Only manuscripts disclosing a funding source other than none (n = 54).

Conclusions The editorial process makes a wide variety of contributions to improve the clarity, accuracy, and reporting of results. Authors often do not disclose potential conflicts of interest or the role of the funding source in submitted manuscripts.

Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA 94118, USA, e-mail: leek@pharmacy.ucsf.edu

Comparison of Author and Editor Suggested Reviewers in Terms of Review Quality, Timeliness, and Recommendation for Publication

Sara Schroter,^{1,2} Leanne Tite,¹ Andrew Hutchings,² and Nick Black²

Objective Many journals give authors the opportunity to suggest reviewers to review their paper. We report a study comparing author-suggested reviewers (ASRs) and editor-suggested reviewers (ESRs) of 10 biomedical journals in a range of specialties to investigate differences in review quality, timeliness, and recommendation for publication.

Design Original research papers sent for external review at 10 participating journals between April 1, 2003, and December 31, 2003, in which the author had suggested at least 1 reviewer were included. Editors were instructed to make decisions about their choice of reviewers in their usual manner. Journal administrators then requested additional reviews from the author's list of suggestions according to a strict protocol using the journals' electronic manuscript tracking systems. Review quality was rated independently using the validated Review Quality Instrument by 2 raters blind to reviewer identity and status. Timeliness was calculated as the interval between dates when reviews were solicited and completed. Recommendation was calculated for 6 journals as proportion rec-

ommending acceptance (including minor revision), resubmission, or rejection. Reviewers who were suggested by both the editor and the author were treated as ASRs.

Results There were 788 reviews for 329 manuscripts. Review quality and timeliness did not differ significantly between ASRs and ESRs (TABLE 3). The ESRs were less likely to provide a recommendation of accept and accept or resubmit. There was no evidence that the effect of reviewer status on review quality, timeliness, and recommendation to accept and recommendation to accept or resubmit varied across journals.

Conclusions Author and editor suggested reviewers of biomedical research in a range of specialties did not differ in the quality of their reviews, but ASRs tended to make more favorable recommendations for publication.

¹BMJ Editorial Office, BMA House, Tavistock Square, London WC1H 9JR, UK, e-mail: sschroter@bmj.com; ²Health Services Research Unit, London School of Hygiene and Tropical Medicine, London, UK

Effect of Authors' Suggestions Concerning Reviewers on Manuscript Acceptance

Lowell A. Goldsmith,¹ Elizabeth Blalock,² Heather Bobkova,² and Russell P. Hall³

Objective To determine the effects of authors' suggested/excluded reviewers on manuscript acceptance.

Design The *Journal of Investigative Dermatology* is ranked first in impact factor in its clinical and basic science specialty. A total of 228 consecutive submissions of original articles, representing one third of 2003 annual submissions, were analyzed for the effect of authors' suggestions concerning reviewers on peer review outcome. Odds ratios (ORs) were calculated with STATA version 7 and are presented with the 95% confidence intervals (CIs).

Results Forty percent of submitters neither suggested nor excluded reviewers; 39% suggested only, 16% both suggested and excluded, and 5% excluded only. Fifty-five percent of submitters suggested between 1 and 14 reviewers (mode, 4). Nineteen percent of suggested reviewers were invited; of those, 92% agreed to review; of the 92%, 90% completed reviews. Twenty-one percent of submitters asked to exclude 1 or more reviewers. Of those requesting exclusion of reviewers, 54% named 1; 25% named 2; and 21% named 3 or more. The most common reason was "close competitors," but usually no reason was given. Authors' requests for exclusion were followed 93% of the time. Odds ratios for acceptance were 1.64 (95% CI, 0.97-2.79) for authors suggesting reviewers compared with those not suggesting reviewers, and 2.38 (95% CI, 1.19-4.72) for those excluding reviewers. Multivariate analysis of acceptance rate had an OR of 2.17 (95% CI, 1.08-4.38) for excluding reviewers, after correcting for the authors suggesting reviewers. No significant difference (*t* test) was found in the time-to-first decision for either group.

Table 3. Impact of Reviewer Status on Review Quality, Timeliness, and Recommendation to Publish

	Author Suggested	Editor Suggested	Difference in Effect Between Journals
Quality			
Mean RQI score	2.58	2.64	ANOVA
Effect of paired differences in scores, mean (95% CI)	1.00	0.05 (-0.04 to 0.15)	<i>P</i> = .74
Timeliness, d			
Median (IQR)	18 (10 to 27)	18 (12 to 26)	Kruskal-Wallis
Wilcoxon signed rank test for paired difference, <i>P</i> value		<i>P</i> = .57	<i>P</i> = .41
Recommendation to publish			
Accept, No. (%)	93 (55.7)	105 (49.5)	
Resubmit, No. (%)	50 (29.9)	56 (26.4)	
Reject, No. (%)	24 (14.4)	51 (24.1)	
Odds ratio (95% CI) from conditional (on manuscript) logistic regression			Likelihood ratio test
Accept (vs resubmit or reject)	1.00	0.61 (0.38 to 0.98)	<i>P</i> = .26
Accept or resubmit (vs reject)	1.00	0.38 (0.20 to 0.70)	<i>P</i> = .47

Abbreviations: ANOVA, analysis of variance; CI, confidence interval; IQR, interquartile range; RQI, Review Quality Instrument.

Conclusions Authors excluding reviewers had higher acceptance rates by univariate analysis and multivariate analysis. Suggesting reviewers had less of an influence on acceptance than excluding reviewers. The advantage for authors of excluding reviewers requires processes to eliminate bias during peer review.

¹Department of Dermatology, University of North Carolina, Chapel Hill, NC, USA; ²*Journal of Investigative Dermatology*, 920B Airport Rd, Suite 216, Chapel Hill, NC 27514, USA, e-mail: elizabeth_blalock@med.unc.edu; ³Dermatology Division, Duke University, Durham, NC, USA

Assessment of Blind Peer Review on Abstract Acceptance for Scientific Meetings

Joseph S. Ross,¹ Cary P. Gross,¹ Yuling Hong,² Augustus O. Grant,³ Stephen R. Daniels,⁴ Vladimir C. Hachinski,⁵ Raymond J. Gibbons,⁶ Timothy J. Gardner,⁷ and Harlan M. Krumholz¹

Objective To determine whether blind peer review, a common method for minimizing reviewer bias, affects acceptance of abstracts to scientific meetings when examined by characteristics of the authors' institutions.

Design We used American Heart Association (AHA) data, which used open peer review from 2000 to 2001 and blind peer review from 2002 to 2004 for abstracts submitted to its annual Scientific Sessions. Authors' institutions were categorized by country (United States vs non-United States) and country's official language (English vs non-English). The US institutions were scored for prestige, combining total National Institutes of Health (NIH) awards (0-2 points) and US News Heart & Heart Surgery hospital ranking (0-2 points), and subsequently categorized as more (3-4 points) or less (0-2 points) prestigious. Data were analyzed using χ^2 tests and logistic regression.

Results On average, the AHA received 13 456 abstracts per year over our study period. A small but significantly lower rate of abstracts were accepted during the blind review period (30% vs 28%, $P < .001$). Blind peer review was associated with lower acceptance from US institutions (41% vs 33%, $P < .001$), while acceptance was slightly but significantly greater from non-US institutions (23% vs 24%, $P < .001$). Among non-US institutions, blind review was associated with slightly greater acceptance from non-English-speaking countries (21% vs 23%, $P < .001$) and trended toward lower acceptance from English-speaking countries (31% vs 29%, $P = .053$). Adjusting for US acceptance rates, blind review was associated with a larger decrease in acceptance from more prestigious US institutions (51% vs 38%) than from less prestigious US institutions (37% vs 32%, $P < .001$), along with lower acceptance from US federal research institutions (65% vs 46%, $P < .001$). There was no difference in acceptance from corporate institutions (29% vs 28%, $P = .79$).

Conclusions Our results suggest there is bias in the open peer review process, favoring prestigious institutions within the United States. These findings argue for universal adoption of blind peer review by scientific meetings.

¹Yale University School of Medicine, IE-61 SHM, PO Box 208088, New Haven, CT 06520-8088, USA, e-mail: joseph.s.ross@yale.edu; ²American Heart Association, Dallas, TX, USA; ³Duke University, Durham, NC, USA; ⁴Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; ⁵University of Western Ontario, London, Ontario, Canada; ⁶Mayo Clinic and Mayo Foundation, Rochester, MN, USA; ⁷Christiana Care Health System, Newark, DE, USA

Peering at Peer Review: Harnessing the Collective Wisdom to Arrive at Funding Decisions About Grant Applications

Nancy E. Mayo,¹ James Brophy,¹ Mark S. Goldberg,¹ Marina B. Klein,¹ Sydney Miller,² Robert Platt,¹ and Judith Ritchie¹

Objective There is a persistent degree of uncertainty and dissatisfaction with the peer review process around the awarding of grants, underlining the need to validate current procedures. The purpose of our study was to compare the allocation of research funding among 3 different processes for peer review: Classic Structured Scientific In-depth 2-reviewer Critique (CLASSIC), all panel members' independent ranking method (RANKING), and committee of review of rankings (RE-RANK).

Design Two consecutive years of a pilot project competition at a major university medical center provided the material for a series of experiments. During the first year of the competition, 11 reviewers rated 32 applications; during the second year, 15 reviewers rated 23 applications. For each of these 2 samples, agreement between the RANKING and CLASSIC methods was compared; the impact of rater on the funding decision was assessed by determining the proportion of projects that would meet the funding cutoff considering all possible pairs of reviewers [$n(n-1)/2$]; Cronbach α was used to identify the number of reviewers needed for optimal consistency, and associations between pairs of raters were calculated.

Results Agreement between the CLASSIC and RANKING methods was poor in both samples ($\kappa = 0.36$). Depending on the pairings, the top rated project in each stream would have failed the funding cutoff with a frequency of 9% and 35%, respectively. Four of the top 10 projects identified by RANKING had a greater than 50% of not being funded by the CLASSIC ranking. Ten reviewers provided optimal consistency for the RANKING method in the first sample but this was not repeated in the second sample. There were only 3 statistically significant positive associations among the 105 pairings of the 15 raters, and there were many negative correlations, some quite strongly negative. Compared with the CLASSIC method, the RANKING resulted in 18 of 23 projects (78%) changing their rank by more than 3 places; however, disagreement on funding status occurred for only 4 of the projects (17%). RE-RANKING during committee discussion resulted in a change of funding status for 2 of the 23 previously ranked projects (<9%).

Conclusions The lack of concordance among reviewers on the relative merits of individual research grants indicates that under the classical, small number-reviewer process, there is a risk that funding outcome will depend on who is assigned as reviewers rather than the merits of the project. Prior RANKING appears to be a way of harnessing the collective wisdom and producing some stability in funding decisions across time.

¹Division of Clinical Epidemiology R4.29, McGill University Health Center, RVH Site, 687 Pine Ave W, Montreal, Quebec, H3A 1A1, Canada, e-mail: nancy.mayo@mcgill.ca; ²Concordia University, Montreal, Quebec, Canada

A Second Order of Peer Review: Peer Review for Clinical Practitioners

Brian Haynes, Chris Cotoi, Leslie Walters, Jennifer Holland, Nancy Wilczynski, Dawn Jedraszewski, James McKinlay, and Ann McKibbin; for the PLUS project

Objective Clinical journals serve many lines of communication, including scientist-to-scientist, clinician-to-scientist, scientist-to-clinician, and clinician-to-clinician. We describe a second order of peer review to serve individual clinician readers.

Design We developed an online clinical peer review system to identify articles of highest interest for each of a broad range of clinical disciplines. Practicing clinicians are recruited to the McMaster Online Rating of Evidence (MORE) system and register according to their clinical discipline. Research staff reviewed more than 110 clinical journals to select each article that meets critical appraisal criteria for diagnosis, treatment, cause, course, and economics of health care problems. An automated system assigns each qualifying article to 4 clinical raters for each pertinent discipline, and records their online assessments of the article's relevance and newsworthiness. Rated articles are transferred to a database that is used to select articles for 3 evidence-based journals and to feed online alerting services: McMaster PLUS and *BMJ*UPDATES+. Users of these services receive alerts and can search the database according to peer ratings for their own discipline. McMaster PLUS is being evaluated in a cluster randomized trial.

Results To date, MORE has 2015 clinical raters, with 34 974 ratings collected for 6573 articles. Ratings for articles of potential interest to both primary care physicians and specialists reflect the different interests of these groups ($P < .05$). Preliminary results from the McMaster PLUS trial ($n = 203$ physicians) show an increase in the use of evidence-based information resources (mean difference, 0.43 logins per month; 95% confidence interval, 0.02-0.16; $P < .02$).

Conclusions A peer review system to suit the information interests of specific groups is clearly feasible. Preliminary findings show that it differentiates between clinical disciplines and stimulates the users of the services it creates to use more evidence-based resources, such as original and systematic review articles.

Health Information Research Unit, Michael G. DeGroote School of Medicine, McMaster University, 1200 Main St W, Room 2C10b, Hamilton, Ontario L8N 3Z5, Canada, e-mail: bhaynes@mcmaster.ca

Scientific Misconduct

Retractions in the Research Literature: Misconduct or Mistakes?

Benjamin G. Druss,¹ Sara Bressi,² and Steven C. Marcus²

Objective While considerable attention has been directed to cases of scientific misconduct in the scientific literature, far less is known about the number and nature of unintentional research errors. To better understand this issue, we examined the characteristics of retracted articles indexed in MEDLINE, focusing on comparisons between misconduct and mistakes.

Design All retractions of publications indexed in MEDLINE between 1982 and 2002 were extracted and categorized by 2 reviewers as representing misconduct (fabrication, falsification, or plagiarism), unintentional error (mistakes in sampling or data analysis, failure to reproduce findings, or omission of information), or other causes. Our analyses compared the characteristics of the retracted article with the timing of the subsequent notification of retraction between articles that were retracted due to unintentional error and scientific misconduct.

Results Of a total of 395 articles retracted during the study period, only 107 (27.1%) reflected scientific misconduct. A total of 244 retractions (61.8%) represented unintentional errors and 44 (11.1%) represented other issues or had no information on the cause of the retraction. Compared with unintentional errors, cases of misconduct were more likely to be written by a single author (10.5% vs 5.7%). Mistakes were more likely to be reported by an author of the initial manuscript (90.2% vs 35.2%), to be in a manuscript with no reported funding source (59.4% vs 40.5%), and to have a shorter time lapse between the initial publication and the retraction (mean, 2 vs 3.3 years). There were no differences in types of retractions based on the date of publication (before or after 1991) or the type of research (human subjects vs basic research).

Conclusions The findings suggest that unintentional mistakes are a substantially more common cause of retractions in the biomedical literature than scientific misconduct. Substantial differences exist between these 2 categories in publication characteristics, authorship, and reporting.

¹Rollins School of Public Health, 1518 Clifton Rd NE, Room 606, Atlanta, GA 30322, USA, e-mail: bdruss@emory.edu; ²University of Pennsylvania School of Social Work, Philadelphia, PA, USA

Citation of Literature Flawed by Scientific Misconduct

A. Victoria Neale, Justin Northrup, Judith Abrams, and Rhonda Dailey

Objective To determine the extent to which authors cite articles identified in official reports as affected by scientific misconduct, and to characterize the nature of such citations.

Design We identified 102 articles from either the National Institutes of Health (NIH) Guide for Grants and Contracts, Findings of Scientific Misconduct, or the US Office of Research Integrity annual reports between 1993 and 2001 as needing retraction or correction, and determined the type of corrigenda posted in PubMed.

Using the Web of Science, we conducted bibliometric analyses to identify subsequent citations of these affected/problem articles. A stratified random sample of 604 was drawn from the population of 5164 citing articles for a content analysis to determine how the affected articles were used by subsequent researchers.

Results Of the 102 articles flawed by misconduct, 4 were not indexed in PubMed; 47 were tagged in PubMed with a retraction notice, 26 with an erratum, and 12 with a comment correction. Ten articles only had a link to the NIH Guide Findings of Scientific Misconduct and 3 articles had no corrigendum whatsoever. Most problem articles were basic science studies (68% in vitro and 7% animal) and 25% were clinical studies. The problem articles had a median of 26 subsequent citations (range, 0-592). The problem article was embedded in a string of references in 61% of citing articles; there was specific reference to the problem article in 39%. Few citing articles (9%) used the problem article as direct support or contrast, 54% used it as indirect support or contrast, and 33% did not address invalid information in the problem article. Only 4% of citing articles referenced the corrigendum.

Conclusions Although most articles named in misconduct investigations have an identifiable corrigenda, few citing articles reference the corrigenda in their bibliography. There is scant evidence of awareness of misconduct in most citing articles.

Department of Family Medicine, Wayne State University, 101 E Alexandrine, Detroit, MI 48201, USA, e-mail: vneale@med.wayne.edu

For Which Cases of Suspected Misconduct Do Editors Seek Advice? An Observational Study of All Cases Submitted to COPE

Sabine Kleinert,¹ Jeremy Theobald,² Elizabeth Wager,³ and Fiona Godlee⁴

Objective To describe the nature and main concerns of all cases discussed at the Committee of Publication Ethics (COPE) from its inception in April 1997 to September 2004.

Design Observational study reporting number of all cases and breakdown of main problems discussed by COPE. Analysis of cases as published in COPE reports and, for 2004, from COPE committee meeting minutes.

Results From 1997 to September 2004, 212 cases were submitted to COPE. The most frequent problem presented was duplicate or redundant publication or submission (58 cases), followed by authorship issues (26 cases), lack of ethics committee approval (25 cases), no or inadequate informed consent (22 cases), falsification or fabrication (19 cases), plagiarism (17 cases), unethical research or clinical malpractice (15 cases), and undeclared conflict of interest (8 cases). Six cases of reviewer misconduct and 3 cases of editorial misconduct were discussed. A total of 132 cases (63%) were relating to papers before publication and 72 (34%) to papers that had already been published; in 8 cases, the question or concern was not related to a particular paper. In the majority of cases (n = 159 [75%]), the committee felt that there were sufficient grounds to pursue cases further. In 79 cases, outcomes were reported following advice by COPE and actions by editors. In 16 cases (20%), authors were exonerated, 15 (19%) could not be resolved satisfactorily, and in 23 (29%), the authors' institution was asked to investigate.

Investigations took longer than 1 year in 36 cases (46%).

Conclusions Editors are confronted with a wide spectrum of suspected research and publication misconduct before and after publication and have a duty to pursue such cases. COPE provides a forum for discussion and advice among editors and a repository of cases for educational purposes. Reports of outcomes suggest that, in many cases, investigations take a long time and a substantial number of cases cannot be adequately resolved.

¹The *Lancet*, 32 Jamestown Rd, London NW1 7BY, UK, e-mail: sabine.kleinert@lanet.com; ²John Wiley & Sons Ltd, London, UK; ³Sideview, Prines Risborough, UK; ⁴*BMJ*, London, UK

SATURDAY, SEPTEMBER 17

Publication Bias and Funding/Sponsorship

Relation of Funding Source, Methodological Quality, and Study Outcome in Bipolar Disorder Drug Reports

Federico Soldani,^{1,2,3} S. Nassir Ghaemi,^{1,2} and Ross J. Baldessarini^{1,3}

Objective Since high quality of clinical research and its reporting are critical for progress in psychiatry and for rational and effective practice, the Clinical Knowledge and Research Methodology Audit (ClinKARMA) Project is systematically evaluating recent research reports in leading psychiatric journals. We report preliminary findings pertinent to bipolar disorder.

Design Published drug treatment studies (2000–2004) were identified by MEDLINE and PsycINFO searches in 5 psychiatric journals with highest impact factors for 2004 as assessed by the Institute for Scientific Information. Funding sources and study outcome were assessed as reported in print. Methodological quality was evaluated according to study design (scored at 5 levels of evidence: double-blind randomized trial, single-blind or nonblind randomized trial, large nonrandomized study, small nonrandomized study, case series or case reports), conclusions based on *P* values (ie, a confounded measure), assessment of precision (confidence intervals or power analysis), drop-out rates in prospective studies, citation of literature supporting research design, and data analyses.

Results Among 75 original articles, 42 of 46 sponsored by pharmaceutical companies were favorable to sponsors' products vs 18 of 29 studies lacking such funding (risk ratio [RR], 1.47; 95% confidence interval [CI], 1.09-1.98), 13 of 46 vs 6 of 29 were randomized (RR, 1.37; 95% CI, 0.58-3.19), and 38% overall relied on baseline to end-point contrasts without controls. Among prospective studies, 21 of 35 vs 14 of 19 had drop-out rates of 40% or greater (RR, 0.81; 95% CI, 0.56-1.19); 14 of 46 vs 10 of 29 had methodological references (RR, 0.88; 95% CI, 0.45-1.72). A minority of studies addressed precision, and virtually all study conclusions relied on *P* values and statistical significance testing.

Conclusions Pharmaceutically funded studies of bipolar disorder tend to have a randomized design and lower drop-out rates but are

less likely to address precision. There is a positive funding-outcome association: explanations may include careful industry selection of which studies get funded, publication or design/analytic bias, and lack of an appropriate comparator or dosage in pharmaceutically sponsored trials. This preliminary literature analysis suggests that, even in highly respected journals, most reports on bipolar disorder are of limited methodological quality and that source of funding may have an important influence on psychopharmacology research results, conclusions, and dissemination.

¹Harvard Medical School, Department of Psychiatry, McLean Division of Massachusetts General Hospital, 115 Mill St, Belmont, MA 02478, USA, e-mail: federico@post.harvard.edu; ²Cambridge Health Alliance, Cambridge, MA, USA; ³McLean Division of Massachusetts General Hospital, Belmont, MA, USA

Are Authors' Financial Ties With Pharmaceutical Companies Associated With Positive Results or Conclusions in Meta-analyses on Antihypertensive Medications?

Veronica Yank,¹ Drummond Rennie,^{2,3} and Lisa A. Bero³

Objective To determine whether authors' financial ties with pharmaceutical companies are associated with positive results or conclusions in meta-analyses on antihypertensive medications.

Design Meta-analyses published January 1966 to June 2002 that evaluated antihypertensive medications in nonpregnant adults were included. We plan to update our search to include meta-analyses published through December 2004. Meta-analyses were identified by electronically searching PubMed and the Cochrane Database and by hand-searching the reference lists of identified meta-analyses. Duplicate meta-analyses were excluded. Non-English-language articles have not yet been evaluated for inclusion. Financial ties were defined as author affiliation and funding source, as disclosed in the article. Results and conclusions were separately categorized as positive (significantly in favor of study drug), negative (significantly against study drug), not significant/neutral, or unclear. The data extraction tool was pretested, and the quality of each meta-analysis was assessed using a validated instrument. A pilot study demonstrated good reliability between the 3 authors in data extraction and quality assessment.

Results Seventy-one eligible meta-analyses were identified. Twenty-three of these (32%) had drug industry financial ties. We found no difference between meta-analyses with or without financial ties in the proportion of positive results. In contrast, conclusions in meta-analyses with financial ties were positive in 91%, unclear in 4%, neutral in 4%, and negative in none, whereas conclusions in meta-analyses without financial ties were positive in 72%, unclear in 2%, neutral in 17%, and negative in 8%. The mean quality scores for each group were similar (0.36 and 0.34 for meta-analyses with and without financial ties, respectively). We will perform multiple logistic regression analyses to determine whether other factors are associated with positive conclusions.

Conclusions Preliminary data suggest that meta-analyses with and without disclosed financial ties to the pharmaceutical industry are similar in direction of results and quality. But those with financial ties have a higher proportion of positive conclusions in favor of the study drug.

¹University of Washington, 4411 4th Ave NE, Seattle, WA 98105, USA, e-mail: vyank@u.washington.edu; ²JAMA, Chicago, IL, USA; ³Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA, USA

Sponsorship, Bias, and Methodology: Cochrane Reviews Compared With Industry-Sponsored Meta-analyses of the Same Drugs

Anders W. Jørgensen and Peter C. Gøtzsche

Objective Trials sponsored by the pharmaceutical industry are often biased. One would therefore expect systematic reviews sponsored by the industry to be biased. We studied whether Cochrane Reviews and industry-sponsored meta-analyses of the same drugs differed in methodological quality and conclusions.

Design We searched MEDLINE, EMBASE, and the Cochrane Library (issue 1, 2003) to identify pairs of meta-analyses—1 Cochrane Review and 1 industry-sponsored review—that compared the same 2 drugs in the same disease and were published within 2 years of each other. Data extraction and quality assessment were performed independently by 2 observers using a validated scale to judge the scientific quality of the reviews and a binary scale to grade conclusions.

Results One hundred seventy-five of 1596 Cochrane Reviews had a meta-analysis that compared 2 drugs. We found 24 paper-based meta-analyses that matched the Cochrane Reviews; 8 were industry-sponsored, 9 had unknown support, and 7 had no support or were supported by nonindustry sources. On a scale from 0 through 7, the overall median quality score was 7 for Cochrane Reviews, 2 for industry-sponsored reviews ($P < .01$), and 2 for reviews with unknown support ($P < .01$). Compared with industry-sponsored reviews, more Cochrane Reviews had stated their search methods, had comprehensive search strategies, used more sources to identify studies, made an effort to avoid bias in the selection of studies, reported criteria for assessing the validity of the studies, used appropriate criteria, and described methods of allocation concealment, excluded patients, and excluded studies. All reviews supported by industry recommended the experimental drug without reservations vs none of the Cochrane Reviews ($P < .001$). Reviews with unknown support and reviews with not-for-profit support or no support also had cautious conclusions.

Conclusions Systematic reviews of drugs should not be sponsored by industry. And if they are, they should not be trusted.

The Nordic Cochrane Centre, Rigshospitalet, Dept 7112, Blegdamsvej 9, DK-2100 Copenhagen, Denmark, e-mail: pcg@cochrane.dk

Is Everything in Health Care Cost-effective? Reported Cost-effectiveness Ratios in Published Studies

Chaim M. Bell,^{1,2,4} David R. Urbach,^{1,3,4} Joel G. Ray,^{1,2}
Ahmed Bayoumi,^{1,2} Allison B. Rosen,⁵ Dan Greenberg,⁶
and Peter J. Neumann⁷

Objective Cost-effectiveness analysis can inform policy makers about the efficient allocation of resources. Incremental cost-effectiveness ratios are used commonly to quantify the value of a diagnostic test or therapy. We investigated whether published studies tend to report favorable cost-effectiveness ratios (less than \$20 000, \$50 000, and \$100 000 per quality-adjusted life year [QALY] gained) and evaluated study characteristics associated with this phenomenon.

Design We reviewed published English-language cost-effectiveness analyses cited in MEDLINE between 1976 and 2001. We included original cost-effectiveness analyses that measured health effects in QALYs. All incremental cost-effectiveness ratios were measured in US dollars set to the year of publication (we did not adjust them to constant dollars because we tested whether ratios targeted certain thresholds, such as \$50 000/QALY, in the year of publication). For all 533 articles, we used generalized estimating equations to determine study characteristics associated with incremental cost-effectiveness ratios below 3 threshold values (\$20 000, \$50 000, and \$100 000/QALY), including journal impact factor in the year prior to publication, disease category, country of origin, funding source, and assigned quality score.

Results About half (712/1433) of all reported incremental cost-effectiveness ratios were less than \$20 000/QALY. Industry-funded studies were more likely to report ratios less than \$20 000/QALY (adjusted odds ratio [OR], 2.2; 95% confidence interval [CI], 1.4-3.4), \$50 000/QALY (OR, 3.5; 95% CI, 2.0-6.1), and \$100 000/QALY (OR, 3.4; 95% CI, 1.6-7.0) than non-industry-funded studies. Conversely, studies of higher methodological quality (OR, 0.58; 95% CI, 0.37-0.91) and those conducted in Europe (OR, 0.59; 95% CI, 0.33-1.1) and the United States (OR, 0.44; 95% CI, 0.26-0.76) were less likely to report incremental cost-effectiveness ratios less than \$20 000/QALY than those conducted elsewhere.

Conclusions The majority of published cost-effectiveness analyses report highly favorable incremental cost-effectiveness ratios. Awareness of potentially influential factors can enable health journal editors as well as policy makers to better judge cost-effectiveness analyses.

¹University of Toronto, Toronto, Ontario, Canada; ²St Michael's Hospital, 30 Bond St, Toronto, Ontario M5B 1W8, Canada, e-mail: belle@smh.toronto.on.ca; ³University Health Network, Toronto, Ontario, Canada; ⁴Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; ⁵University of Michigan, Ann Arbor, MI, USA; ⁶Health Systems Management, Ben-Gurion University of the Negev, Beersheba, Israel; ⁷Center for Risk Analysis, Harvard School of Public Health, Cambridge, MA, USA

Publication Bias and Journal Factors

Characteristics of Accepted and Rejected Manuscripts at Major Biomedical Journals: Predictors of Publication

Kirby P. Lee, Elizabeth A. Boyd, Peter Bacchetti, and Lisa A. Bero

Objective To identify characteristics of submitted manuscripts that are associated with acceptance for publication at major biomedical journals.

Design Prospective cohort of original research articles (n = 1107) submitted for publication January 2003–April 2003 at 3 leading biomedical journals. Studies of experimental and observational design, systematic reviews, and qualitative, ethnographic, or nonhuman studies were included. Case reports of single patients were excluded. Characteristics identified for each manuscript included research focus, study design, analytic methods (statistical/quantitative or descriptive/qualitative), clearly stated hypothesis, statistical significance of primary outcome ($P < .05$ or not $P > .05$), sample size (dichotomized at median, n = 73), description of participants, funding source, country of origin and institutional affiliation of corresponding author, and sex and academic degree of authors. Multivariate logistic regression was used to model predictors of publication.

Results Of 1107 manuscripts submitted, 68 (6.1%) were published. Characteristics associated with publication are listed in TABLE 4. When added to the model shown, a statistically significant primary outcome did not appear to improve chance of publication (significant vs not; odds ratio, 0.55; 95% confidence interval, 0.24–1.30; $P = .18$).

Table 4. Characteristics Associated With Manuscripts Published in 3 Leading Biomedical Journals

Characteristic	Category	OR (95% CI)	P Value
Study design	RCT (reference)	1.00	
	Non-RCT	0.40 (0.19-0.86)	.02
	Systematic review	0.79 (0.09-6.97)	.83
	Other	NA	NA
Analysis	Descriptive vs statistical	3.17 (1.69-5.95)	<.001
Sample size	>73 vs <73	2.62 (1.19-5.80)	.02
Funding source	Any disclosure vs no disclosure	2.06 (1.10-3.84)	.02

Abbreviations: CI, confidence interval; NA, not available because numbers too small for analysis; OR, odds ratio; RCT, randomized controlled trial.

Conclusions We found no evidence of editorial bias favoring studies with statistically significant results. Submitted manuscripts with randomized, controlled study design or systematic reviews, analyses using descriptive or qualitative methods, sample sizes of 73 or more, and disclosure of the funding source were more likely to be published.

Institute for Health Policy Studies, University of California, San Francisco, 3-333 California St, Suite 420, Box 0613, San Francisco, CA 94118, USA, email: leek@pharmacy.ucsf.edu

Rethinking Publication Bias: Developing a Schema for Classifying Editorial Discussion

Kay Dickersin and Catherine Mansell

Objective To identify new features of publication bias (ie, favorable factors associated with decision to publish) in editorial decision making.

Design Cross-sectional, qualitative analysis of discussions at manuscript meetings of *JAMA*. One of us (K.D.) attended 12 editorial meetings in 2003 and took notes recording discussion surrounding 102 manuscripts. In addition, editors attending the meetings noted the “negative” (not favoring publication) and “positive” (favoring publication) factors associated with each manuscript considered. We extracted unique sentences and phrases used by the editors to describe the manuscripts reviewed and entered them into an Excel spreadsheet; one of us (C.M.) subsequently coded the phrases using NVivo2 qualitative analysis software.

Results From the list of phrases, we sorted terms such as author characteristics (eg, highly regarded), the manuscript (eg, well-written), and the topic, as well as peer reviewer opinion and scientific aspects of the work (eg, sample size). We determined that most discussion could be categorized into 3 broad categories: scientific features, journalistic goals, and writing. We revised our classification system iteratively, using internal discussion and peer input until we established 27 subgroups. Phrases related to scientific merit (design, measures, population) predominated. Phrases relating to the journalistic goals (readership needs, timeliness) were nearly as common, followed by phrases related to the writing. Although statistical significance of study findings was rarely explicitly discussed, terms related to the concept of strength and direction of findings were relatively frequent.

Conclusions Studies of editorial publication decision making should assess the relative impact of factors related to scientific merit, journalistic goals, and writing, in addition to statistical significance of results.

Brown University, Center for Clinical Trials and Evidence-based Healthcare, 169 Angell St, Box G-S2, Providence, RI 02912, USA, e-mail: kay_dickersin@brown.edu

Is Publication Bias Associated With Journal Impact Factor?

Yuan-I Min,¹ Aynur Unalp-Arida,² Roberta Scherer,³ and Kay Dickersin⁴

Objective To evaluate the association between statistically significant results for the primary outcome and the journal impact factor of published clinical trials.

Design An existing data set from a retrospective follow-up study (1988-1989) of 3 cohorts of initiated studies evaluating publication bias was used for this analysis. Only trials with full publication (defined as an article with ≥ 3 pages) after study enrollment had been completed, at the time of the follow-up, were included ($n = 219$). Results for primary outcomes and publication history of the trials were based on investigator interviews. Journal impact factor of published trials was assessed using the 1988 Science Citation

Index (impact factors of 0 were assigned to journals not in the Science Citation Index). We used the journal impact factor of the first full publication after study enrollment was completed for all analyses. If more than 1 article was published in the same year, the best journal impact factor was used. The association between journal impact factor and statistical significance of the primary results was examined using the odds ratio.

Results Journal impact factors in our cohorts of trials ranged from 0 to 21.148 (published in 1967-1989). The numbers of journals in each journal impact factor quartile were 45 (first quartile), 35 (second quartile), 26 (third quartile), and 13 (fourth quartile). The median journal impact factors for trials with statistically significant primary results ($n = 133$) and those with nonsignificant primary results ($n = 86$) were 2.73 and 1.91, respectively ($P = .02$). Compared with the lowest journal impact factor quartile, the odds ratios of having a significant primary result, for each journal impact factor quartile from low to high (≤ 0.672 , > 0.672 and ≤ 2.241 , > 2.241 and ≤ 4.482 , > 4.482), were 1.00, 1.24 (95% confidence interval [CI], 0.57-2.70), 1.39 (95% CI, 0.66-2.93), and 2.54 (95% CI, 1.14-5.63), respectively ($P = .02$ for trend).

Conclusions Trials with statistically significant results were more likely to be published in a high-impact journal. Our results suggest publication bias at the journal level. Data from other cohorts of trials are needed to generalize our results.

¹MedStar Research Institute, Department of Epidemiology and Statistics, 6495 New Hampshire Ave, Suite 201, Hyattsville, MD, USA, e-mail: nancy.min@medstar.net; ²Johns Hopkins University, Center for Clinical Trials, Baltimore, MD, USA; ³University of Maryland, Department of Epidemiology and Preventive Medicine, Baltimore, MD, USA; ⁴Brown University, Department of Community Health, Providence, RI, USA

Effect of Indexing, Open Access, and Journal “Phenomena” on Submissions, Citations, and Impact Factors

Impact of SciELO and MEDLINE Indexing on the Submission of Articles to a Non-English-Language Journal

Danilo Blank, Claudia Buchweitz, and Renato S. Procionoy

Objective To evaluate the impact of Scientific Library Online (SciELO) and MEDLINE indexing on the number of articles submitted to *Jornal de Pediatria*, a Brazilian bimonthly pediatrics journal with a Portuguese print version and a bilingual (Portuguese/English) free-access, full-text, online version.

Design Analysis of total article submission, submission of articles from countries other than Brazil, and acceptance data from 2000 through 2004. Since there were no changes in the editorial board or in the methods of manuscript submission during this period, 3 events were considered as having a potential impact on submission rates: launch of the bilingual Web site, indexing in the SciELO, and MEDLINE indexing. Thus, data analysis was divided into 4 stages: stage I, pre-Web site (January 2000-March 2001 [15 months]); stage II, Web site (April 2001-July 2002 [16 months]); stage III, SciELO (August 2002-August 2003 [13 months]); and stage IV, MEDLINE (September 2003-December 2004 [16 months]). Simple regression was used for trend analysis, 1-way ANOVA was used on rank-transformed data with the Duncan post hoc test to compare the

number of submissions in each period, and the Fisher exact test with Finner-Bonferroni P value adjustment was used to compare submissions from countries other than Brazil in the 4 periods.

Results There was a significant trend toward linear increase in the number of submissions along the study period ($P = .009$). The numbers of manuscripts submitted in stages I through IV were 184, 240, 297, and 482, respectively. The numbers of submissions were similar in stages I and II ($P = .15$) but statistically higher in stage III ($P < .001$ vs stage I and $P = .006$ vs stage II) and stage IV ($P < .001$ vs stages I and II and $P < .05$ vs stage III). The variation in mean monthly submissions became more pronounced in stages III and IV (TABLE 5). The rate of article acceptance decreased during the study period. The number of original articles published has been stable since the 2001 March/April issue ($n = 10$), when the journal reached a printed page limit, leading to stricter judgment criteria and a relative decrease in acceptance rate. The numbers of manuscript submissions from countries other than Brazil in stages I through IV were 1, 2, 0, and 17, respectively, with $P < .001$ for the comparison of stage IV with previous stages.

Table 5. *Jornal de Pediatria*: Mean Monthly Submissions per Analysis Stage, 2000 through 2004

Stage	Mean (95% CI)	SD	SE
I (before Web site)	12.27 (10.24-14.29)	3.654	0.943
II (after Web site)	15.00 (13.52-16.48)	2.781	0.695
III (SciELO)	22.85 (18.92-26.78)	6.504	1.804
IV (MEDLINE)	30.13 (25.38-34.87)	8.899	2.225

Abbreviations: CI, confidence interval; SciELO, Scientific Library Online.

Conclusions These results suggest that SciELO and MEDLINE indexing contributed to increased manuscript submission to *Jornal de Pediatria*. SciELO indexing was associated with an increase in Brazilian submissions, whereas MEDLINE indexing led to an increase in both Brazilian and international submissions.

Jornal de Pediatria, Brazilian Society of Pediatrics, Rua Gen Jacinto Osorio 150/201, Porto Alegre, RS, CEP 90040-290, Brazil, e-mail: blank@ufrgs.br

Effect of Open Access on Citation Rates for a Small Biomedical Journal

Dev Kumar R. Sahu, Nithya J. Gogtay, and Sandeep B. Bavdekar

Objective Articles published in print journals with limited circulation are cited less frequently than those printed in journals with larger circulations. Open access has been shown to improve citation rates in the fields of physics, mathematics, and astronomy. The impact of open access on smaller biomedical journals has not been studied. We assessed the influence of open access on citation rates for the *Journal of Postgraduate Medicine*, a small, multidisciplinary journal that adopted open access without article submission or article access fee.

Design The full text of articles published since 1990 were made available online in 2001. Citations for these articles as retrieved using Web of Science, SCOPUS, and Google Scholar were divided into 2 groups: the pre-open-access period (1990-2000) and the post-open-access period (2001-2004). Citations for the articles published between 1990 through 1999 during these 2 periods were

compared using the unpaired t test or Mann-Whitney test.

Results Of the 553 articles published between 1990 through 1999, 327 articles received 893 citations between 1990 and 2004 (TABLE 6). The 4-year, post-open-access period accounted for 549 (61.5%) of these citations, and 164 articles (50.1%) received their first citations only after open access was provided in 2001. For every volume studied (1990 through 1999), the maximum number of citations per year was received after 2001. None of the articles published during 1990 through 1999 received any citation in the year of publication. In contrast, articles published in 2002, 2003, and 2004 received 3, 7, and 22 citations, respectively, in the year of publication itself.

Table 6. Citations to Articles Published Before and After Open Access to the Small Biomedical Journal Was Provided in 2001

	Pre-Open Access (n = 344)	Post-Open Access (n = 549)	95% CI	P Value
Citations per volume per year, mean (SD)	4.6 (4.4)	13.6 (4.1)	4.8-12.6	<.001*
Citations per 100 articles per year, mean (SD)	5.8 (4.8)	27.0 (8.2)	14.4-26.1	<.001*
Citations per cited article, median (IQR)	0 (0-1)	1 (1-2)	0.0001-1.0	<.001†
Publication to first citation period, median (IQR), y	4 (2.25-5)	1 (1-2)‡	2.2-2.9	<.001†

Abbreviations: CI, confidence interval; IQR, interquartile range.

*By unpaired t test.

†By Mann-Whitney test.

‡Time from open access to first citation for articles first cited after 2001

Conclusions Open access was associated with an increase in the number of citations received by the articles. It also decreased the lag time between publication and the first citation. For smaller biomedical journals, open access could be one of the means for improving visibility and thus citation rates.

Journal of Postgraduate Medicine, 12, Manisha Plaza M. N. Rd, Kurla (W). Mumbai 400070, India, e-mail: dksahu@vsnl.com

More Than A Decade in the Life of the Impact Factor

Mabel Chew,¹ Martin Van Der Weyden,¹ and Elmer V. Villanueva²

Objective To analyze trends in the impact factor of 7 general medical journals (*Ann Intern Med*, *BMJ*, *CMAJ*, *JAMA*, *Lancet*, *Med J Aust* and *N Engl J Med*) over 11 years and to ascertain the views of these journals' past and present Editors-in-Chief regarding major influences on their journal's impact factor.

Design Retrospective analysis of impact factor data from Institute for Scientific Information (ISI) Journal Citation Reports, Science Edition, 1994 to 2004, and interviews with the 10 editors-in-chief of these journals (except *Med J Aust*) who had served between 1999 and 2004, conducted from November 2004 to February 2005.

Results Impact factors generally rose over the 11-year period. However, the relative changes in impact factors (averaged over 11 years) ranged from 6% to 172% per year. The impact factor is calculated yearly by dividing the number of citations that year to any article published by that journal in the previous 2 years (the numerator) by the number of eligible articles published by

that journal in the previous 2 years (the denominator). In general, the numerator for most journals tended to rise over this period while the denominators tended to drop. When standardized for the denominator, the average relative changes fell naturally into 2 groups: those that remained constant (ie, increases in crude impact factors were due largely to rises in numerators) and those that did not (ie, increase in crude impact factors were due largely to falls in denominators) (TABLE 7). Nine of 10 editors-in-chief were contactable, and all agreed to be interviewed. Possible reasons given for rises in citation counts included active recruitment of “high-impact” articles by courting researchers, the inclusion of more journals in the ISI database, and increases in citations in articles. Boosting the journal’s media profile was also thought to attract first-class authors and, therefore, “citable” articles. Many believed that going on-line had not made a difference to citations. Most had no deliberate policy to publish fewer articles (thus altering the impact factor denominator), which was sometimes the unintended result of redesign, publication of longer research articles (ie, fewer could be “fit” in each issue) or of editors being “choosier.” However, 2 editors did have such a policy, as they realized impact factors were important to authors. Concerns about the accuracy of ISI counting for the impact factor denominator prompted a few editors to routinely check their impact factor data with ISI. All had mixed feelings about using impact factors to measure journal quality, particularly in academic culture, and mentioned the tension between aiming to improve impact factors and “keeping their constituents [clinicians] happy.”

Table 7. Average Annual Relative Changes (%) in Impact Factors for 7 Journals

Journal	Crude change (95% CI)	Change standardized for number of published articles (95% CI)
<i>Ann Intern Med</i>	21 (2, 41)	-19 (-51, 14)
<i>BMJ</i>	32 (22, 42)	26 (19, 34)
<i>CMAJ</i>	41 (25, 56)	28 (9, 48)
<i>JAMA</i>	172 (134, 211)	121 (21, 222)
<i>Lancet</i>	9 (-73, 92)	34 (4, 64)
<i>Med J Aust</i>	6 (3, 10)	5 (2, 7)
<i>N Engl J Med</i>	139 (106, 173)	150 (61, 238)

Abbreviation: CI, confidence interval.

Conclusions Impact factors of the journals studied rose in the 11-year period due to rising numerators and/or falling denominators, to varying extents. This phenomenon was perceived by editors-in-chief to occur for various reasons, sometimes including editorial policy. However, all considered the impact factor a mixed blessing—attractive to researchers but not the best measure of clinical impact.

¹*Medical Journal of Australia*, Locked Bag 3030, Strawberry Hills, NSW 2012, Australia, e-mail: mabel@ampco.com.au; ²National Breast Cancer Centre, Camperdown, NSW 2050, Australia

Quality and Influence of Citations and Dissemination of Scientific Information to the Public

Use and Persistence of Internet Citations in Scientific Publications: An Automated 5-Year Case Study of Dermatology Journals

Kathryn R. Johnson,¹ Jonathan D. Wren,² Lauren F. Heilig,¹ Eric J. Hester,¹ David M. Crockett¹, Lisa M. Schilling,¹ Jennifer M. Myers,³ Shayla Orton Francis,⁴ and Robert P. Dellavalle^{1,5}

Objective To examine Internet address and uniform resource locator (URL) citation characteristics in medical subspecialty literature using dermatology journals as a case study.

Design An automated computer program systematically extracted and analyzed Internet addresses cited in the 3 dermatology journals rated highest for scientific impact: *Journal of Investigative Dermatology*, *Archives of Dermatology*, and *Journal of the American Academy of Dermatology*, published between January 1999 and September 2004. Corresponding authors of articles with inaccessible URLs were contacted regarding the content and importance of the inaccessible information.

Results Overall, 7337 journal articles contained 1113 URL citations, of which 18% were inaccessible. The percentage of articles containing at least 1 Internet citation increased from 2.3% in 1999 to 13.5% in 2004. Internet citation inactivity also increased significantly with time since publication, from 11% of those published in 2004 to 35% in 1999 ($P < .001$). The URL inaccessibility was highest in the *Journal of Investigative Dermatology* (22%) and lowest in *Archives of Dermatology* (15%) ($P = .03$). *Archives of Dermatology* was the only journal of the 3 examined with a stated Internet referencing policy in the instructions for authors. For all years, URL accessibility was significantly associated with top-level domain and directory depth but not associated with the presence of an accession date or a tilde in the URL. Of the 204 inaccessible URLs, at least some content was recoverable for 59% via the Internet Archive (<http://www.archive.org>). Thirty-nine of 100 randomly selected citations included accession dates. Results from surveying authors regarding inaccessible URL content and importance will be reported.

Conclusions Internet citations increasingly used and lost in dermatology journals likely reflect URL use and loss across all medical subspecialty literature. Policy changes are needed to stem the loss of cited Internet information in scientific publications.

¹Department of Dermatology, University of Colorado at Denver and Health Sciences Center (UCDHSC), 1055 Clermont St, Box 165, Aurora, CO 80045, USA, e-mail: robert.dellavalle@uchsc.edu; ²University of Oklahoma, Norman, OK, USA; ³Louisiana State University School of Medicine, New Orleans, LA, USA; ⁴UCDHSC School of Medicine, Aurora, CO, USA; ⁵Dermatology Service, Department of Veterans Affairs Medical Center, Denver, CO, USA

Are Reviewers Influenced by Citations of Their Own Work? Evidence From the *International Journal of Epidemiology*

Matthias Egger,^{1,2} Lesley Wood,¹ Erik von Elm,² Anthony Wood,¹ Yoav Ben Shlomo,¹ and Margaret May¹

Objective To examine whether reviewers' assessments of manuscripts are influenced by citations to their own work.

Design The *International Journal of Epidemiology* receives about 700 manuscripts each year, of which about 30% are sent out to reviewers. Reviewers are asked to make a decision using 4 categories (accept as is, minor revision, major revision, and reject). We examined whether the reviewers' decision was influenced by the number of citations to their work or other manuscript characteristics by abstracting data using a standardized proforma. Data were analyzed with logistic regression, accounting for the clustered nature of the data. We included manuscripts refereed from 2003 to February 2005. The study is ongoing; the prespecified sample size is 800 manuscripts.

Results We analyzed 1641 reviewer reports on 637 manuscripts. The median number of reviewers per manuscript was 2.6 (interquartile range [IQR], 2-3) and the median number of citations in manuscripts was 28 (IQR, 22-39). Work of 606 reviewers (37%) was cited. Among manuscripts with at least 1 citation the median number of citations to papers authored by a reviewer was 1.8 (IQR, 1-2; range, 1-9). TABLE 8 shows the reviewers' verdicts on manuscripts according to citations to their own work. The odds of rejecting a paper were reduced when the manuscript cited the reviewer's work: odds ratios compared with manuscripts not citing the reviewer were 0.87 (95% confidence interval, 0.65-1.18) for manuscripts citing 1 paper and 0.74 (95% confidence interval, 0.49-1.11) for manuscripts citing 2 or more papers ($P = .11$ for trend).

Table 8. Reviewers' Verdicts on Manuscripts According to Citations to Their Own Work: Analysis of 1641 Reviewer Reports on 637 Manuscripts Submitted to the *International Journal of Epidemiology*

Reviewer's Verdict	No. (%)			
	Reviewer Not Cited	1 Citation	≥2 Citations	Total
Accept as is	65 (6.3)	23 (5.8)	15 (7.3)	103 (6.3)
Minor revisions	447 (43.2)	184 (46.0)	97 (47.1)	728 (44.4)
Major revisions	304 (29.4)	117 (29.3)	60 (29.1)	481 (29.3)
Reject	219 (21.2)	76 (19.0)	34 (16.5)	329 (20.1)

Conclusions There may be truth in the old professor's adage that you should cite the work of likely reviewers in your papers. Alternative explanations include chance and confounding by other factors. The final analysis, which will be based on a larger number of manuscripts and reviewer reports, will provide more robust evidence.

¹Editorial Office, *International Journal of Epidemiology*, Department of Social Medicine, University of Bristol, UK; ²Department of Social and Preventive Medicine, University of Berne, Finkenhubelweg 11, CH-3012, Berne, Switzerland, e-mail: egger@ispm.unibe.ch

How the News Media Report on Research Presented at Scientific Meetings: More Caution Needed

Steven Woloshin and Lisa M. Schwartz

Objective Scientific meeting presentations garner new media attention, despite the fact that the underlying research is often preliminary and has undergone limited peer review. We examined whether media stories report basic study facts and caveats, specifically regarding the preliminary nature of the research.

Design Three physicians with clinical epidemiology training analyzed front-page newspaper stories ($n = 34$), other newspaper stories ($n = 140$), and television/radio stories ($n = 13$) identified in a Lexis-Nexis/ProQuest search for research reports from 5 scientific meetings in 2003 (American Heart Association, 12th World AIDS Conference, American Society of Clinical Oncology, Society for Neuroscience, and the Radiological Society of North America).

Results Basic facts were often missing: 34% of the 187 stories did not mention study size, 18% did not mention study design (when mentioned, 41% were ambiguous enough that expert readers could not be certain about the design), and 40% did not quantify the main result (when quantified, 35% presented relative change statistics without a base rate—a format known to exaggerate the perceived magnitude of findings). Important study caveats were often missing. Among the 124 stories covering studies other than randomized trials, 85% failed to mention (explicitly mention or imply) any relevant study limitation (eg, imprecision of small studies, no comparison group in case series, confounding in observational studies). Among the 61 stories reporting on animal studies, case series, or studies with fewer than 30 people, 67% did not highlight their limited relevance to human health. While 12 stories mentioned a corresponding “in press” medical journal article, only 3 of the remaining 175 noted that the findings were either unpublished, might not have undergone peer review, or might change as the study matured.

Conclusions News stories about scientific meeting research presentations often omit basic study facts and caveats. Consequently, the public may be misled about the validity and relevance of the science presented.

VA Outcomes Group (111B), Department of Veterans Affairs Medical Center, White River Junction, VT 05009, USA, e-mail: steven.woloshin@dartmouth.edu

Reading Between the Ads: Assessing the Quality of Health Articles in Top Magazines

Brad Hussey,¹ Diane Miller,² and Alejandro R. Jadad³

Objective This research examines the quality of health-related reporting published in the top 100 US consumer magazines, using a validated index of the quality of written consumer health information.

Design The authors purchased and reviewed consumer publications published in August 2003 and likely to include health-related information from the Audit Bureau of Circulations top 100 magazine rankings based on average paid circulation figures. Articles were included if they made overt claims about the effect of interventions

of any kind on health outcomes. Disagreements on article inclusion were settled by consensus. The first eligible article in each section of each publication was reviewed to ensure consistency. Articles were evaluated individually by all 3 authors using 7 questions of the DISCERN evaluation tool that address factual information. Each question was answered in terms of whether the article met or did not meet specific criteria. Consensus was reached for all answers.

Results Thirty-two magazines, with a combined number of more than 100 million subscribers, met the inclusion criteria and 57 articles were deemed eligible for review. Only 2% of the articles dated the source of information quoted, 7% provided references for each claim made, and 14% gave additional contact information for sources. Most articles failed to acknowledge uncertainty about the participant interventions (74%), risks associated with the interventions (74%), or other possible treatment choices (56%).

Conclusions Health articles as found in top consumer magazines fail to meet basic quality criteria and could mislead the public. Publishers of consumer magazines are missing an opportunity to increase the current low levels of functional health literacy of the public, at little or no cost. There is an opportunity for peer-reviewed biomedical publications to provide support to, and help define standards for, the lay press, thereby protecting the public.

¹Dundas, Ontario, Canada; ²Ontario Hospital Association, Toronto, Ontario, Canada; ³Centre for Global eHealth Innovation, University Health Network and University of Toronto, R. Fraser Elliott Building, 4th Floor, 190 Elizabeth St, Toronto, Ontario M5G 2C4, Canada, e-mail: ajadad@uhnres.utoronto.ca

SUNDAY, SEPTEMBER 18

Reporting of Studies: Abstracts and Publication After Meeting Presentations

Trials Reported in Abstracts: The Need for a Mini-CONSORT

Sally Hopewell and Mike Clarke

Objective To assess the need for a better reporting standard (such as a mini-CONSORT) for trials reported in abstracts.

Design A total of 209 randomized trials were identified from the proceeding of the American Society of Clinical Oncology conference in 1992. Full publications were identified for 125 trials by searching The Cochrane Central Register of Controlled Trials and PubMed (median time to publication 27 months; interquartile range, 15-43). A sample of 40 trials was selected within a specific area of cancer; 4 trials were later excluded. A checklist (based on CONSORT) was used to compare information reported in the 36 conference abstracts with that reported in the corresponding abstract of the subsequent full publications. Steps were taken to blind the source of each abstract.

Results Some aspects of trials were well-reported in the conference abstract and full publication abstract: 92% (n = 33) of study Objectives, 94% (n = 34) of participant eligibility, 100% (n = 36) of trial interventions, and 89% (n = 32) of primary outcomes were the same in both abstracts (TABLE 9). Other areas were more discrepant: 92% (n = 33) of conference abstracts and 97% (n = 35) of full publication abstracts reported number of participants randomized; but these numbers were the same only 42% (n = 15) of the time. Likewise, 69% (n = 25) of conference abstracts and 44% (n = 16) of full publication abstracts reported number of participants analyzed, and these numbers were the same only 11% (n = 4) of

Table 9. Reporting Criteria Assessed for Trials Reported in Abstracts

Criteria Assessed		No. (%)		
		Conference Abstract	Paper Abstract	Level of Agreement*
Objectives	Study objectives were described	35 (97)	34 (94)	33 (92)
	Trial name or identifier given	13 (36)	3 (8)	3 (8)
	Date of trial given	20 (56)	9 (25)	7 (19)
Study quality	Method of allocation concealment was described	0 (0)	1 (3)	0 (0)
	Method of blinding was described	6 (17)	6 (17)	6 (17)
Participants	Characteristics of eligible participants were described	34 (94)	35 (97)	34 (94)
Interventions	Experimental intervention was described	36 (100)	36 (100)	36 (100)
	Comparator intervention was described	36 (100)	36 (100)	36 (100)
	Participants randomized to experimental intervention were described	25 (69)	18 (50)	6 (17)
	Participants randomized to comparator intervention were described	25 (69)	18 (50)	6 (17)
Outcomes	Primary outcome measure was described	35 (97)	35 (97)	32 (89)
	Trial status was described (eg, closed)	6 (17)	20 (56)	6 (17)
	No. of participants randomized was described	33 (92)	35 (97)	15 (42)
	No. of participants analyzed was described	25 (69)	16 (44)	4 (11)
	Intention-to-treat principle was described	4 (11)	3 (8)	4 (11)
	Important adverse effects were described	25 (69)	25 (69)	24 (67)
Results	Results for primary outcome were described	36 (100)	32 (89)	4 (11)
Conclusions	Primary conclusions were described	34 (94)	36 (100)	29 (80)

*Percentage of agreement between what was reported in the conference abstract and in the abstract of the full publication.

the time. This is partly because the conference abstracts were often preliminary reports. Lack of information was a major problem in assessing trial quality; 0% of conference abstracts and 3% (n=1) of full publication abstracts reported method of allocation concealment, 11% (n = 4) of conference abstracts and 8% (n = 3) of full publication abstracts reported intention-to-treat analyses.

Conclusions Previous research has shown that trials presented as conference abstracts are poorly reported. However, this study suggests that they may contain as much or more useful information than the abstract of a full publication. The quality of reporting of trials in abstracts, both in the proceedings of scientific meetings and in journals, needs to be improved because this may be the only accessible information to which someone appraising a trial has access.

The UK Cochrane Centre, Summertown Pavilion, Middle Way, Oxford, OX2 7LG, UK, shopewell@cochrane.co.uk

Are Relative Risks and Odds Ratios in Abstracts Believable?

Peter C. Gøtzsche

Objective An abstract should give the reader a quick and reliable overview of the study. I studied the prevalence of significant *P* values in abstracts of randomized controlled trials and observational studies that reported relative risks or odds ratios and checked whether reported *P* values in the interval .04 to .06 were correct.

Design I searched MEDLINE and included all abstracts of articles published in 2003 that contained the words “relative risk” or “odds ratio” and appeared to report results from a randomized controlled trial (N=260). For comparison, random samples of 130 abstracts from cohort studies and 130 abstracts from case-control studies were selected. I noted the *P* value for the first relative risk or odds ratio that was mentioned, or calculated it if it was not reported.

Results The first result in the abstract was reported to be statistically significant in 70% of randomized controlled trials, 84% of cohort studies, and 84% of case-control studies (TABLE 10). Many of these significant results were derived from subgroup or secondary analyses. *P* values were more extreme in observational studies ($P < .001$), and more extreme in cohort studies than in case-control studies ($P = .04$, Mann-Whitney test). The distribution of *P* values around $P = .05$ was skewed for randomized controlled trial; *P* values ranged from .04 to .05 for 29 trials, and from .05 to .06 for only 5 trials. I could verify the calculations for 23 and 4 of these trials, respectively. The 4 nonsignificant results were all correct, whereas 17 of the 23 significant results (74%) appeared to be wrong (11 of the 17 were easy to verify, because the authors used Fisher exact test, and 6 were highly doubtful). Recalculation was rarely possible for observational studies because the presented results had been adjusted for confounders.

Table 10. Reporting of *P* Values in Abstracts of Studies

<i>P</i> Value	No. (%)		
	Randomized Controlled Trials	Cohort Studies	Case-Control Studies
<.0001	24 (9)	38 (29)	20 (15)
<.001	25 (10)	16 (12)	14 (11)
<.01	40 (15)	27 (21)	31 (24)
<.02	20 (8)	7 (5)	16 (12)
<.03	18 (7)	6 (5)	11 (8)
<.04	25 (10)	7 (5)	11 (8)
<.05	29 (11)	8 (6)	6 (5)
<.10	16 (6)	3 (2)	4 (3)
<.20	10 (4)	4 (3)	5 (4)
≥.20	53 (20)	14 (11)	12 (9)
Total	260	130	130

Conclusions Significant results in abstracts are exceedingly common but are often misleading because of erroneous calculations and probably also repetitive trawling of the data.

The Nordic Cochrane Centre, Rigshospitalet, Department 7112, Blegdamsvej 9, DK-2100 Copenhagen Ø, Denmark, e-mail: pcg@cochrane.dk

Do Clinical Trials Get Published After Presentation at Biomedical Meetings? A Systematic Review of Follow-up Studies

Erik von Elm¹ and Roberta Scherer²

Objective To determine the rate at which abstracts describing results of randomized or controlled clinical trials are subsequently published in full and to investigate study factors associated with full publication.

Design Systematic review, searching MEDLINE, Embase, and the Cochrane Library from the start of each database until May 2005 for reports of bibliographic follow-up studies examining the subsequent full publication of randomized or controlled clinical trial results initially presented in abstract or summary form. We identified additional reports through citations (Science Citation Index), and reference lists of included reports, author files, and by verbal communication. We excluded reports with less than 2 years of follow-up after the meeting or without information on numbers of abstracts examined or followed. We calculated weighted mean rates of full publication. Time to publication was analyzed using survival analysis. Dichotomous data for study factors were analyzed using relative risk and a random effects model.

Results In 23 reports (published 1990-2004) there were follow-up data of 3946 controlled clinical trial abstracts from a wide range of biomedical disciplines. Follow-up from the meeting ranged from 29 to 300 months. Ultimately, 2351 abstracts were fully published; the weighted mean rate of full publication was 57.5% (95% confidence interval, 55.9-59.2). Using survival analysis the estimated publication rate was 62.9% (95% confidence interval, 60.7-65.2) at 9 years. Study factors associated with publication included positive results, higher sample size, industry funding, higher quality of abstract reporting, and multicenter status (TABLE 11).

Table 11. Pooled Estimates for Associations Between Study Factors Described in Meeting Abstracts and Subsequent Full Publication

Study Factors	No. of Included Follow-up Studies	No. of Abstracts Followed	Relative Risk for Subsequent Full Publication (95% Confidence Interval)
Positive results vs not positive results	10	1285/1362	1.17 (1.09–1.27)
Higher sample size vs lower sample size	7	808/747	1.21 (1.09–1.34)
Industry funding vs other funding	2	119/268	1.38 (1.18–1.62)
Higher-quality reporting vs lower-quality reporting	2	77/155	1.30 (1.00–1.71)
Multicenter status vs single-center status	5	582/626	1.28 (1.06–1.56)
Oral presentation vs poster presentation	4	397/432	1.31 (0.97–1.76)
Reporting of harms, yes vs no	1	81/139	1.09 (0.84–1.40)

Conclusions Results from almost half of all randomized and controlled clinical trials initially presented at biomedical meetings were not made accessible to researchers through full publication. Furthermore, we found evidence for publication bias in the controlled clinical trials that investigators did publish in full. For systematic reviews on efficacy and harm of treatment, authors cannot rely solely on results from fully published controlled clinical trials.

¹Department of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11 CH-3012 Bern, Switzerland, e-mail: vonelm@ispm.unibe.ch; ²Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

Improving the Quality of Reporting of Trials and Other Studies

Design, Analysis, and Presentation of Crossover Trials

Edward J. Mills,¹ An-wen Chan,² Ping Wu,³ Gordon H. Guyatt,¹ and Douglas G. Altman⁴

Objective Although crossover trials enjoy wide use, standards for analysis and reporting have not been established. We reviewed methodological aspects and quality of reporting in a representative sample of published crossover trials.

Design We searched MEDLINE for December 2000 and identified all randomized crossover trials. We abstracted data independently, in duplicate, on 14 design criteria, 13 analysis criteria, and 14 criteria assessing the data presentation.

Results We identified 526 randomized controlled trials, of which 116 were crossover trials. Trials were drug efficacy (49%), pharmacokinetic (21%), and nonpharmacologic (30%). The median sample size was 15 (range 6–361). Most (65%) trials used 2 treatments and had 2 periods (64%). Few trials reported allocation concealment (15%; 95% confidence interval [CI], 10%–22%) or

sequence generation (7%; 95% CI, 7%–13%). Only 14% of trials reported a sample size calculation and only 38% of these considered pairing of data in the calculation. Carry-over issues were addressed in 30% (95% CI, 23%–39%) of trial's methods, yet only 11.5% (95% CI, 8%–20%) explicitly tested for carry-over effects. Most trials reported and defended a washout period (68.7%, 95% CI, 62%–78%). Almost all trials (96%; 95% CI, 91%–98%) tested for treatment effects using paired data and also presented details on by-group results (95%; 95% CI, 89%–97%). Only 28% (95% CI, 21%–36%) presented CIs or SE so that data could be entered into a meta-analysis. Just 3% (95% CI, 1%–8%) of trials provided a participant flow diagram and 44% (95% CI, 37%–54%) of trials did not account for all participants in the analysis. Seven trials (5%; 95% CI, 3%–11%) analyzed the first period as a separate trial.

Conclusions Reports of crossover trials frequently omit important methodological issues in design, analysis, and presentation. Guidelines for the conduct and reporting of crossover trials might improve the conduct and reporting of studies using this important trial design.

¹Department of Clinical Epidemiology & Biostatistics, McMaster University, HSC-2C12, 1200 Main St West, Hamilton, Ontario, Canada L8N 3Z5, e-mail: millsej@mcmaster.ca; ²Department of Medicine, University of Toronto, Toronto, Ontario, Canada; ³Department of Epidemiology, London School of Hygiene and Tropical Medicine, London, UK; ⁴Centre for Statistics in Medicine, Oxford, UK

Quality of Trials in Operative Surgery: Where Is the Comic Opera?

Catherine Jane Walter,¹ Walter Jo Dumville,² Catherine Hewitt,² David Torgerson,² Philip Drew,¹ and John R. T. Monson¹

Objective Surgical randomized controlled trials (RCTs) have previously been criticized for using weak methodology. This study aims to investigate their quality.

Design All surgical RCTs published in the *BMJ*, *JAMA*, *Lancet*, and *New England Journal of Medicine* from 1998 to 2004 were identified. For each surgical RCT included, a nonsurgical RCT, matched for journal and publication year, was also randomly selected. In each paper adequacy of the randomization sequence generation, allocation concealment, power calculation, and recruitment were assessed using predefined criteria. These aspects of trial design are equally achievable and reportable in all trials regardless of their medical speciality. Differences in adequacy between the surgical and nonsurgical RCTs were compared using the t-test (adjusted for intracluster correlation).

Results There was no significant difference in the quality of reporting between surgical and nonsurgical trials (TABLE 12). When compared, adequate reporting of randomization sequence generation was seen in 42% of surgical trials and 30% of nonsurgical trials and adequate allocation concealment was recorded in 46% and 47%, respectively. When combining these 2 interrelated steps of randomization together, just 26% of surgical trials and 23% of nonsurgical trials reported both adequately. Adequate recruitment was recorded in 52% of surgical and 55% of nonsurgical trials with approximately a quarter of the trials in both the surgical and nonsurgical categories reporting an adequate power calculation. The ratio of adequate vs inadequate power calculations amongst

all the trials with adequate recruitment was approximately 1:2, this difference was not statistically significant ($P = .58$; 95% confidence interval, -40.1 to 24.7).

Table 12. Methodology of Reports of Surgical RCTs vs Nonsurgical RCTs

Methodology Examined	No. (%)		P Value	95% Confidence Interval
	Surgical RCTs (n = 66)	Non-Surgical RCTs (n = 66)		
Adequate randomization sequence generation	28 (42.4)	20 (30.3)	.40	-20.4 to 44.6
Adequate allocation concealment	30 (45.5)	31 (47.0)	.89	-29.8 to 26.6
Adequate randomization sequence generation and allocation concealment together	17 (25.8)	15 (22.7)	.81	-26.8 to 32.9
Adequate power calculation	17 (25.8)	16 (24.2)	.80	-12.7 to 15.8
(Interim analysis excluded)	n = 63	n = 60		
Adequate recruitment	33 (52.4)	33 (55.0)		
Adequate recruitment with adequate power calculation	10 (30.3)	11 (33.3)	.79	-30.4 to 24.3

Abbreviation: RCTs, randomized controlled trials.

Conclusions The quality of surgical trials is no different than nonsurgical trials with no statistically significant differences seen between these 2 groups. However, approximately half or less of all the trials reviewed reported adequate methodology. Better execution and reporting of all health care trials are required.

¹Academic Surgical Unit, The University of Hull, Castle Hill Hospital, Cottingham, Hull HU16 5JQ, UK, e-mail: c.j.walter@hull.ac.uk; ²The University of York, Heslington, York, UK

Does the CONSORT Checklist Improve the Quality of Reports of Randomized Controlled Trials? A Systematic Review

Amy C. Plint,^{1,2} David Moher,^{1,2} Kenneth Schulz,³ Douglas G. Altman,⁴ and Andra Morrison⁵

Objective To systematically review whether the adoption of the CONSORT checklist is associated with improvement in the quality of reporting of randomized controlled trials.

Design We searched MEDLINE, Embase, Cochrane Central, and reference lists of included studies. Studies were eligible if they: (1) compared CONSORT adopters and non-adopters after the publication of the CONSORT statement, (2) compared CONSORT adopting journals before and after publication of the CONSORT statement, or (3) compared CONSORT adopters and non-adopters before and after the publication of the CONSORT statement. We imposed no restriction on publication language. Two reviewers independently determined eligibility and resolved disagreements by consensus. The journal's instructions to authors were reviewed to determine if they had adopted the CONSORT statement at the time of publication of the RCT. Outcomes examined included reports for any of the 22 items on the CONSORT checklist or overall trial quality. Data was extracted by a single reviewer and

checked for accuracy by a second reviewer. A checklist based on principles of internal and external validity was used to evaluate study quality.

Results A total of 1131 studies were retrieved, of which 250 were considered possibly relevant. Seven studies were included in the final analysis. Most studies were quasi-experimental studies with no control group. Five studies compared CONSORT adopters and non-adopters after the publication of the CONSORT statement, 1 study compared CONSORT adopting journals before and after publication of the CONSORT statement, and 1 study compared CONSORT adopters and non-CONSORT adopters before and after publication of the CONSORT statement. Overall 34 outcomes were examined with individual studies reporting between 1 and 19 outcomes. Such variability resulted in most outcomes being reported in only 1 study. Within this limitation, CONSORT adopters had significantly better reporting of the method of sequence generation, allocation concealment, and overall number of CONSORT items than non-CONSORT adopters. Reporting of participant flow was however no different between these 2 groups. In studies examining CONSORT-adopting journals before and after the publication of the CONSORT statement, description of the method of sequence generation, participant flow and total CONSORT items were improved after adoption by the journal of the CONSORT statement but description of allocation concealment was not. The influence of the manner in which a journal adopts the CONSORT statement on randomized controlled trial reporting will also be reported (for example, requiring submission of a completed checklist with manuscript vs requiring only that studies be "reported according to the CONSORT statement").

Conclusions This study was limited by the lack of similar outcome reporting between studies. Overall it appears that journal adoption of CONSORT results in better description of the method of randomization and overall reporting of more CONSORT checklist items but may have variable effects of reporting of allocation concealment and participant flow.

¹University of Ottawa, 401 Smyth Rd, Ottawa, Ontario, Canada K1H 8L1, e-mail: plint@cheo.on.caottawa; ²Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada; ³Family Health International, Research Triangle Park, NC, USA; ⁴Cancer Research UK Medical Statistics Group, Centre for Statistics in Medicine, Oxford, UK; ⁵Canadian Coordinating Office for Health Technology Assessment, Ottawa, Ontario, Canada

Does the STARD Statement Improve the Quality of Reporting of Diagnostic Accuracy Studies?

Nynke Smidt,¹ Anne W. S. Rutjes,² Daniëlle A. W. M. van der Windt,¹ Raymond W. J. G. Ostelo,¹ Johannes B. Reitsma,² Patrick M. Bossuyt,² Lex M. Bouter,¹ and Henrica C. W. de Vet¹

Objective To determine whether the publication of the STARD statement has improved the quality of reporting of diagnostic accuracy studies in journals with an impact factor of at least 4. The STARD statement consists of 25 items and encourages the use of a flow diagram.

Design The quality of reporting in articles of primary studies of diagnostic accuracy published in 2000 (pre-STARD) and 2004 (post-STARD) is compared in journals adopting the STARD

statement (*Annals of Internal Medicine*, *BMJ*, *Clinical Chemistry*, *JAMA*, *Lancet*, *Neurology*, and *Radiology*) and nonadopting journals (*Archives of Internal Medicine*, *Archives of Neurology*, *Circulation*, *Gut*, and *New England Journal of Medicine*). Two reviewers independently evaluated the quality of reporting of included articles, using the STARD statement. A total STARD score for each article was calculated by summing up the number of reported items. Differences in mean STARD scores between studies published in 2000 and 2004 were analyzed using random coefficient analyses, taking journal level effects into account.

Results A total of 124 articles published in 2000 and 141 articles published in 2004 were included. The mean STARD score (range 0-25 points) was 11.9, (range 3.5-19.5) for studies published in 2000, and 13.6 (range 4.0-21.0) for studies published in 2004. After adjustment for journal, the mean difference (95% confidence interval) of the STARD score between studies published in 2000 and 2004 is 1.81 (95% confidence interval, 0.61-3.01). The mean improvement of the quality of reporting between studies published in 2000 and 2004 was not influenced by the design of the study (cohort or case-control) or by whether or not journals had adopted the STARD statement. Studies published in 2004 reported significantly more often methods for calculating test reproducibility of the index test (16% vs 35%), the distribution of the severity of disease in those with the target condition and other diagnoses (22% vs 52%), estimates of variability of diagnostic accuracy between subgroups (39% vs 60%), and a flow diagram (2% vs 12%).

Conclusions Although the quality of reporting in articles on diagnostic accuracy is improved after the publication of the STARD statement, the quality of reporting is still less than optimal. Authors, editors, and reviewers should pay more attention to the reporting of articles by reviewing the items of the STARD statement and including a flow diagram.

¹Institute for Research in Extramural Medicine, VU University Medical Center, Van der Boechorststraat 7, 1081 BT, Amsterdam, the Netherlands, e-mail: n.smidt@vumc.nl; ²Department of Clinical Epidemiology & Biostatistics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands

Quality of Reporting Trials: Protocols, Manuscripts, Published Articles, and Postpublication

Comparison of Submitted and Published Reports of Randomized Trials

Douglas G. Altman,¹ John P. A. Ioannidis,² David Moher,³ Jill Mollison,¹ David L. Schriger,⁴ and Sara Schroter⁵

Context Published articles describing randomized controlled trials have reported widespread deficiencies in both study methodology and reporting. Important discrepancies have been found between trial protocols and details given in subsequently published journal articles. It is unknown whether the publication process increases or decreases such deficiencies and discrepancies.

Objective To characterize differences between reports of randomized controlled trials as originally submitted to the *BMJ* and as published.

Design For a cohort of 75 consecutive randomized controlled trial reports submitted to *BMJ* in 2001, we will compare the original manuscript with the subsequently published article. We will assess both versions of each article and document the extent and nature of differences. We will relate any changes to referees' comments from the submission to the *BMJ*. Both versions of the same trial report will be assessed by 2 people independently. We will collect detailed information about changes, such as in reported study methods or presentation of results and figures. We will contact authors for those articles for which no publication was found to discover reasons for nonpublication.

Results So far, 59 of 75 articles have been published in the *BMJ* (n = 12) or elsewhere (n = 47). We will describe differences between the submitted and published articles with respect to: (1) title, authorship, and length; (2) adherence to key elements of the CONSORT statement, especially details of randomization, blinding, sample size calculation, and specification of outcomes; (3) numerical results and style of presentation, considering in particular the consistency of statements about primary outcomes; (4) reporting of harms; and (5) authors' conclusions. We will also characterize differences in methodology and findings between studies that were and were not published, and describe reasons given by authors for non-publication.

Conclusions This study will describe the impact of the publication process on the quality of reporting and content of reports of randomized controlled trials.

¹Cancer Research UK/NHS Centre for Statistics in Medicine, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, UK, e-mail: doug.altman@cancer.org.uk; ²Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; ³Chalmers Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; ⁴UCLA Emergency Medicine Center, UCLA School of Medicine, Los Angeles, CA, USA; ⁵*BMJ* Editorial Office, BMA House, London, UK

Reporting of Study Outcomes in Comparative Drug Trials: Evidence From Proposals Submitted to a Research Ethics Committee and Corresponding Full Publications

Erik von Elm,¹ Karin Huwiler,¹ Mark Witschi,¹ Alexandra Röllin,¹ Charles Senessie,¹ Nicola Low,¹ and Matthias Egger^{1,2}

Objective To study the reporting of results from comparative drug trials in a comprehensive set of study proposals and corresponding publications.

Design Ongoing longitudinal study based on all 1681 proposals submitted to the Ethics Committee of the Medical Faculty, University of Berne, Switzerland, from 1988 to 1998. We classify the study design of all proposals and identify comparative drug trials. Corresponding publications were searched in Cochrane Central and retrieved. From both proposals and publications, 2 investigators independently extract the number and content of outcomes and assess their concordance. We will extract information on statistical significance of outcomes stated in protocols that are or are not reported in publications. We plan to examine the reasons for nonpublication of studies and outcomes by contacting applicants.

Results A random sample of 1137 (68%) proposals have been assessed to date. The largest group were comparative drug trials (418 of 1137, 37%). Other designs were comparative trials of other interventions (239, 21%), noncomparative studies of interventions (299, 26%), observational studies (143, 13%), and laboratory studies (31, 3%). Seven proposals (1%) were not classifiable. We now have performed literature searches for 95 of 418 comparative drug trials. For 46 of 95 (48%) proposals, we identified 117 corresponding publications published between 1994 and 2001. A pilot random sample of 8 (of 46) proposals and 12 (of 117) corresponding publications were assessed for outcome concordance. Of 93 outcomes specified in 8 protocols, 26 (28%) were not reported in any of the corresponding 12 publications. Of 76 outcomes reported in publications, 9 (12%) were not specified in the corresponding protocols (TABLE 13). Analyses of a larger data set will be presented.

Table 13. Reporting of Outcomes in a Pilot Sample of Proposals of Comparative Drug Trials

Outcomes		Reported in Publications		Total
		Yes	No	
Specified in proposals	Yes	67	26	93
	No	9	n/a	9
Total		76	26	102

Conclusions Preliminary analyses indicate that a substantial proportion of comparative drug trials were not published at least 6 years after submission of proposals to a research ethics committee. Underreporting occurs not only on the level of entire studies but also of single outcomes. Prespecified outcomes are omitted in subsequent publications and, conversely, outcomes are reported that had not been specified in the proposals.

¹Department of Social and Preventive Medicine, University of Berne, Finkenhubelweg 11, CH-3012 Berne, Switzerland, e-mail: egger@ispm.unibe.ch; ²Department of Social Medicine, University of Bristol, Bristol, UK

A Comparison of the Published Version of Randomized Controlled Trials in a Specialist Clinical Journal With the Original Trial Protocols

Lucy Chappell,¹ Zarko Alfirevich,¹ Patrick Chien,¹ Susan Jarvis,² and Jim G. Thornton^{1,2}

Objective To measure the degree to which changes in design between protocol and publication may compromise the validity of randomized controlled trial results.

Design A survey of randomized controlled trials published in *BJOG: An International Journal of Obstetrics and Gynaecology* over the 4-year period 2001 to 2004. Authors and research ethics committees were contacted to provide copies of the trial protocol.

Results Results are available for 2 of the 4 years (the full results will be reported). For these 2 years, 53 randomized trials were identified. Nine authors could not be contacted, and the 44 who responded supplied protocols for 30 trials. Three were not in English and 1 was for the wrong trial, leaving 26 for analysis. In 4 trials, the published description of the intervention differed significantly from that in the protocol. In 6 trials, the published sample size was smaller than that planned in the protocol, and in 11 it was larger. A reason for the difference was given in only 2 trials. In only 6 trials was the published primary outcome exactly the same as that in the protocol, and in only 9 trials was the analysis method stated in both the protocol and published version.

Conclusions It is disappointing that we were unable to obtain trial protocols from 23 of 53 authors. Among the remainder, important changes between protocol and published paper are common. These may seriously compromise the validity of published research. Journal editors should encourage researchers to register their protocols in a public forum or compare protocols and papers before publication. If alterations between protocol and published version are unavoidable, readers should be made aware of them.

¹*BJOG: An International Journal of Obstetrics and Gynaecology*; ²University of Nottingham, Hucknall Road, Nottingham NG5 1PB, UK, e-mail: jim.thornton@nottingham.ac.uk

Trial Bank Publishing of Randomized Trials: Preliminary Results

Ida Sim and Ben Olatov

Objective Publishing randomized clinical trials (RCTs) into machine-understandable “trial banks” may allow computers to better help clinicians practice evidence-based medicine. The RCT Bank trial bank can capture more than 160 aspects of trial design, execution, and results. Our objective was to partner with journals to co-publish RCTs into the RCT Bank.

Design We invited authors of RCTs published by *JAMA* or *Annals of Internal Medicine* between January 2002 and July 2003 to co-publish their trial in the RCT Bank. We entered all information from participating manuscripts and from authors themselves, where necessary, into the RCT Bank using a secure Web-based tool. Completed entries were released under open access at RCT Presenter (rctbank.ucsf.edu/Presenter/). We conducted an online survey of clinicians and Cochrane Collaboration and US Evidence-Based Practice Center systematic reviewers comparing RCT Presenter reports to the corresponding journal articles.

Results During the project period, 54 of 108 RCTs met inclusion criteria. The author participation rate rose from 38% to 76% after an example of a co-published trial was available. Fourteen diverse RCTs were co-published, covering a variety of clinical domains, intervention types (eg, drugs, procedures), outcome types (eg, categorical, survival), and result types (eg, efficacy analysis). We also captured methodological information such as blinding and follow-up. Data entry time averaged 1 to 2 hours while extracting information from the manuscript averaged 6 to 8 hours. A total of 30 of 83 survey respondents provided satisfaction results. These 30 respondents rated RCT Presenter better than journal articles on speed (85%), ease of use (81%), information organization (85%), and clarity (65%), and better or comparable to journal articles on

trial understanding (73%) and use in clinical care (85%). The only feature where a majority of respondents favored the journal article over RCT Presenter was trustworthiness (64%). Fifty percent preferred using RCT Presenter over journal articles. Overall, 70% of respondents rated RCT Presenter as good as or better than journal articles for all surveyed attributes.

Conclusions We have demonstrated proof of concept and user satisfaction with trial bank publishing.

Department of Medicine and Program in Biological and Medical Informatics, University of California, San Francisco, 3333 California St, Suite 435 Q, San Francisco, CA 94143-1211, USA, e-mail: sim@medicine.ucsf.edu

Hot Topics: Open Access and Trial Registries

Authors' Access to Financial Support at the Time of Paper Acceptance: A Survey of Biomedical Journals

Sara Schroter, Leanne Tite, and Ahmed Kassem

Objective The success of open access publishing funded through author charges (author pays model) is dependent on authors having access to financial support at the time of submission or acceptance of their research papers. Our objective was to determine the availability of external funding for author publication charges at different stages in the research process.

Design We conducted a cross-sectional electronic survey of the corresponding authors of all research articles (including reviews) published in *BMJ*, *Archives of Disease in Childhood*, and *Journal of Medical Genetics* in 2003.

Results A total of 377 of 524 (72%) authors responded to the survey and 62% (233 of 377) said they received external funding to support their study but with notable differences between journals. The majority of research (64%, 150 of 233) was funded with some public money, 36% (83 of 233) with some charity money, and 7% (16 of 233) with some industry money. Among the respondents, 2.1% (8 of 377) were funded by the National Institutes of Health, 2.4% (9 of 377) by Wellcome Trust, and 5% (19 of 377) by MRC (England), whom already included or planned to build publication fees into their grants. Of the funded authors, 44% (102 of 233) had access to funding at the time of manuscript submission and 41% (95 of 233) at the time of acceptance. The average duration of funding was 28.3 (SD 24) months, but 31% (72 of 233) received funding for 1 year or less. Non-externally funded research was largely supported through departmental resources (56%, 80 of 144) and/or by carrying out research in own time (63%, 91 of 144). Overall, 29% (109 of 377) of authors reported that they spend all or most of their time working on nonfunded research.

Conclusions A large proportion of published research is not externally funded, and many funded researchers currently do not have access to financial support at the time their paper is

accepted for publication. Under current conditions, the viability of an open access, author-pays business model is questionable, and may prohibit the publication of some research.

BMJ Editorial, BMA House, Tavistock Square, London WC1H 9JR, UK, e-mail: sschroter@bmj.com

Would You Drop Your Membership? Professional Organization Members' Reaction to Open Access

Harold C. Sox and Wayne Bylsma

Objective Sponsoring organizations are concerned about the financial impact of completely free access to the entire contents of the web version of a journal. We surveyed members of a professional organization to assess the potential impact of open access on membership renewals.

Design We sent a self-administered mail survey to a stratified sample of US, non-student, current, below aged 65 American College of Physicians (ACP) members. We received 1008 valid responses (52% response rate). One question asked if electronic access were free to the public at the time of publication, would the respondent continue to be an ACP member? We weighted response rates to be proportional to the ACP's US non-student member population, aged 65 and younger in terms of age groups (35, 36-55, 56-65 y), race (white vs nonwhite), specialty (general internal medicine vs others), and sex. We based 95% confidence intervals on the SE of the estimate.

Results Of the respondents, 80.8% (n = 808) of members answered yes, they would continue to be an ACP member; 3.5% (n = 36; 95% CI, 2.5%-4.9%) said that they would not continue to be a member. The remaining 15.7% (n = 154; 95% CI, 13.5%-18.2%) said that they were uncertain. Of the survey respondents, 36.4% (n = 315) were in the exploratory or getting established phase of their careers; 25.8% (n = 252) said that they accessed journal articles on the Internet a few times per week, 16.1% (n = 150) did so daily, and 12.1% (n = 126) never did. According to survey respondents, *Annals of Internal Medicine* was, by a considerable margin, the most valued of all member benefits.

Conclusions These data suggest that a small proportion, but nonetheless substantial number of members may not continue to be ACP members if we provided open electronic access to *Annals of Internal Medicine*. An even larger minority were uncertain about remaining members. Open access is a potential threat to membership organizations' ability to retain members.

Annals of Internal Medicine and American College of Physicians, Philadelphia, PA 19106, USA, e-mail: hsox@mail.acponline.org

Assessing the Quality of Information Recorded on Trial Registries

Munira Nurbhai,¹ Pasquale Lorenzo Moja,² Jeremy Grimshaw,^{1,3} Alessandro Liberati,² An-Wen Chan,⁴ Kay Dickersin,⁵ Karmela Krleza-Jeric,⁶ David Moher,⁷ Ivan Moschetti,² Drummond Rennie,^{8,9} Ida Sim,⁹ and Jimmy Volmink¹⁰

Objective In 2004 ICMJE member journals agreed to adopt a trial registrations policy as 1 solution to publication bias. The value of trial registries has been well-established and is supported by funding and government agencies. Yet, there is little focus on the type and completeness of data in existing trial registries. Our objectives were to determine the type of data recorded in trial registries, assess their quality, compare structure and content of different trial registries, and identify optimal data structures for them.

Design In January 2005, a random sample of trial records was extracted from 3 international trial registries: ClinicalTrials.gov, US National Cancer Institute (NCI) PDQ Database, and the International Standard Randomised Controlled Trial Number (ISRCTN) Register. No restrictions were placed on trial status, design, or medical area, although some trial registries adopted inclusion criteria. Records were assessed using a quality checklist consisting of 30 desirable items grouped into 5 categories, identifying information, trial details, funding, contact details and data collection forms.

Results A pilot study assessing 45 records (15 from each register) was undertaken. The overall quality was poor (median, 16; range

10-22 items completed per record) and varied considerably across trial registries. ClinicalTrials.gov (median 17) and NCIPDQ (median 18) fared better than ISRCTN (median 13) (ClinicalTrials.gov vs NCIPDQ $P = .79$; ClinicalTrials.gov vs ISRCTN $P < .001$; NCIPDQ vs ISRCTN $P < .001$). Records did not provide complete data in the category of trial details (items: primary outcome; key trial dates) and contact details (items: full address; fax). A copy of the study data collection form was not included in any record. Results from the definitive study will be available for presentation.

Conclusions Although our quality criteria were not stringent, the results show considerable variation in the data available in current trial registries and an unsatisfactory quality of desirable requirements for trial protocols. The results suggest that there is room for standardization of the type of data recorded in trial registries to provide better and more complete essential public information.

¹Clinical Epidemiology Program, Ottawa Health Research Institute, ASB Box 693, Room 2006, 1053 Carling Ave, Ottawa, Ontario K1Y 4E9, Canada, e-mail: mnurbhai@ohri.ca; ²Centro Cochrane Italiano, Istituto Mario Negri, Italy; ³Institute for Best Practices, Institute of Population Health, University of Ottawa, Ottawa, Ontario, Canada; ⁴Dept of Medicine, University of Toronto, Toronto, Ontario, Canada; ⁵Department of Community Health, Brown University, Providence, RI, USA; ⁶Canadian Institutes of Health Research, Ottawa, Ontario, Canada; ⁷Thomas C. Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; ⁸JAMA, Chicago, IL, USA; ⁹University of California, San Francisco, CA, USA; ¹⁰Primary Health Care Directorate, Faculty of Health Sciences, University of Cape Town, South Africa

Notes

Poster Session Abstracts

Poster sessions will be held in the Gold Room.

SATURDAY, SEPTEMBER 17

Authorship and Contributorship

Quantification of Authors' Contributions and Eligibility for Authorship: A Randomized Trial

Ana Ivaniš, Darko Hren, Aleksandra Mišak, Dario Sambunjak, Matko Marušić, and Ana Marušić

Objective To quantify authors' contributions in manuscripts sent to the *Croatian Medical Journal* and their eligibility for authorship criteria defined by the International Committee of Medical Journal Editors (ICMJE) by (1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published.

Design All 532 authors who submitted 121 manuscripts to the *Croatian Medical Journal* from January 18 to May 23, 2005, were randomly allocated into 2 groups and asked to rate their contributions from 0 (none) to 4 (full). A total of 240 were assigned to the rating-questionnaire group (57 manuscripts) and 292 were assigned to the yes/no-questionnaire group (64 manuscripts) for 11 contribution categories. They were not instructed on ICMJE authorship criteria.

Results The number of manuscripts with authors not fulfilling ICMJE criteria was significantly higher in yes/no-questionnaire group than in the rating-questionnaire group (82.8% vs 28.1%; $\chi^2_1 = 36.87$, $P < .001$). The number of undeserved authors was significantly higher in yes/no-questionnaire group than in the rating-questionnaire group (64.7% vs 13.8%; $\chi^2_1 = 140.77$, $P < .001$). The rating-questionnaire group also reported more contribution categories.

Conclusion The lower prevalence of undeserved authorship in the rating-questionnaire group indicates that quantification of authors' contributions may be a more accurate method of contribution disclosure.

Croatian Medical Journal, Zagreb University School of Medicine, Salata 3, HR-10000 Zagreb, Croatia, e-mail: aivanis@mef.hr

Conflict of Interest

Conflict of Interest Policies at Scientific Journals: A Cross-Disciplinary Comparison

Jessica S. Ancker¹ and Annette Flanagin²

Objective To determine the prevalence and types of conflict of interest policies at journals in a variety of scientific disciplines.

Design We targeted the 7 journals with the highest impact factors in each of 12 scientific categories (TABLE 14) from the Institute for Scientific Information's Journal Citation Reports. We examined their published policies in a publication review (August-October 2004) and asked additional questions in an online survey (May-July 2005).

Results We identified 28 policies (33% of the journals) in instructions for authors or editorials. Publicly available policies were available at all 7 general medical journals and 6 chemistry journals but uncommon or absent in other disciplines. Of journals ranked 1 or 2 by impact factor in their category, 54% (13 of 24) had policies; of journals ranked 3-7, 25% (15 of 60) had policies. All policies required disclosure; 2 also imposed bans. Of the 28 policies, 15 (54%) provided no definition of conflict of interest, 22 (79%) did not specify whether disclosures would be used in the review process, and 13 (46%) did not specify whether disclosures could be published. Representatives from 47 journals (56%) completed the survey (TABLE 14). Journals with any type of policy were more likely to report a recent history of problems with financial (13 of 38) and non-financial (14 of 38) conflicts than were journals without any policies (0 of 9 for financial, 2 of 9 for non-financial). For authors, financial disclosure was required in 21 of 23 policies (91%); in cases of financial conflict, authors were barred from publishing in 4 (17%). For reviewers, disclosure of financial conflicts was required in 18 of 22 (82%) policies; in cases of financial conflict, reviewers were requested not to review in 17 (77%) and barred from reviewing in 10 (45%). For editors, financial disclosure was required in 18 of 22 policies (82%); editors were barred from editing articles in cases of financial conflict by 14 (64%) and required to avoid all financial conflicts by 1 (5%).

Table 14. Survey Participation Rates and Frequency of Reported Conflict of Interest Policies

ISI Category	Participation Rate, No. (%)	No. With Policy (Types of Policies)		
		For Authors	For Reviewers	For Editors
Medicine, general and internal	7 (100)	7 (1f, 6b)	5 (5b)	6 (1f, 5b)
Multidisciplinary sciences	7 (100)	5 (1f, 4b)	4 (1nf, 3b)	4 (3b, 1na)*
Psychology	6 (86)	3 (2b, 1na)*	3 (1nf, 2b)	2 (1nf, 1b)*
Plant sciences	5 (71)	3 (2b, 1na)	3 (3b)	3 (3b)
Astronomy-astrophysics	4 (57)	1 (1b)	2 (2nf)	2 (2b)
Medicine, research-experimental	4 (57)	4 (2f, 1nf, 1b)	1 (1b)*	2 (2b)*
Biological sciences	3 (43)	2 (1nf, 1b)	1 (1nf)	2 (2nf)
Engineering-multidisciplinary	3 (43)	1 (1nf)	1 (1nf)	0
Physics	3 (43)	0	0	0
Zoology	3 (43)	1 (1f)	2 (2nf)	1 (1nf)
Biology	2 (29)	1 (1b)	0	0
Chemistry	0			
All	47 (56)	28 (5f, 3nf, 18b, 2na)*	22 (8nf, 14b)*	22 (1f, 4nf, 16b, 1na)*

Abbreviations: ISI, Institute for Scientific Information; f, financial only; nf, nonfinancial only; b, both; na, no answer. Percentages have been rounded.

*Additional policies in development in this category.

Conclusions In this sample of journals from 12 scientific disciplines, conflict of interest policies are common in general medicine and multidisciplinary science but less common or absent in other scientific disciplines. The relative emphasis on financial and non-financial concerns varies across disciplines. Definitions of conflict of interest and other details are lacking in many of the policies.

¹Doctoral Program, Department of Sociomedical Sciences, Mailman School of Public Health, Columbia University, 722 W 168th St, Room 1115, New York, NY 10032, USA e-mail: jsa2002@columbia.edu; ²JAMA, Chicago, IL, USA

Conflicts of Interest Disclosed by Authors of Manuscripts Submitted to a General Medical Journal

Christine Laine, Mary Beth Schaeffer, and Catharine Stack

Objective Many researchers have financial conflicts of interest that could bias their work, but incomplete disclosure policies have made documentation of the frequency and types of conflicts difficult. We describe conflicts that authors disclosed when submitting manuscripts to a general medical journal and examine the association of conflicts with editorial decisions.

Design Beginning August 17, 2004, the journal's electronic manuscript processing system began to require completion of conflict of interest disclosure forms before submission. The forms require reporting the presence or absence of potential conflicts of interest related to institutional-industry relationships and author-industry relationships including employment, consulting, grants, royalties, honoraria, patents, stock ownership, and other. Conflict disclosures are available to all involved in the editorial process. The final analysis will include manuscripts submitted through April 17, 2005. In addition to describing the frequency and types of conflicts and their association with editorial decisions, final analysis will examine factors associated with the presence of conflicts.

Results Preliminary results include 635 manuscripts submitted electronically through December 31, 2004. Of these, 106 (16.7%) had >1 and 9.8% had >2 disclosed conflicts. Over one-third of manuscripts with conflicts had institutional conflicts. The most common types of author conflict were grants (69.8%, 74 of 106), consulting (49.5%, 52), and honoraria (41.9%, 44). Less frequent conflicts related to employment (29.2%, 31), stock ownership/other (20.9%, 22), royalties (10.5%, 11), or patents (16.0%, 17). Preliminary analysis does not show that manuscripts with conflicts (78 rejected of 106, 73.6%) are more likely to be rejected than manuscripts without conflicts (426 of 529, 80.5%).

Conclusions Preliminary results suggest that conflicts are present for a minority of manuscripts submitted to biomedical journals. Almost one-third of disclosed conflicts relate to institutional relationships with industry. The most frequent types of author

conflicts relate to grants, consulting, and honoraria. Manuscripts with conflicts did not have higher rejection rates than manuscripts without conflicts.

Annals of Internal Medicine, 190 N Independence Mall West, Philadelphia, PA, 19106 USA, e-mail: claine@acponline.org

Ethical Concerns

Imaging and Nonimaging Journal Policies Regarding Institutional Review Board Approval and Informed Consent Declarations by Authors

Andrew Y. Choi,¹ Douglas S. Katz,^{1,2} and Anthony V. Proto²

Objective To determine the current policies of imaging journals (ie, journals that focus on radiology topics) and nonimaging journals (ie, clinically oriented medical journals that do not have medical imaging as their main focus) regarding author declaration of institutional review board (IRB) approval and informed consent for prospective and retrospective human research submitted for publication.

Design A survey was mailed to the editorial offices of 63 imaging journals and 104 nonimaging journals. Questions included: "Does your journal have an official policy regarding IRB approval and informed consent for human research studies?"; "Do you require authors to declare if IRB approval and informed consent were obtained for prospective studies, and if IRB approval or exemption was obtained along with informed consent or waiver of informed consent for retrospective studies?"; and "Do you require authors to submit proof of IRB approval?"

Results There were 30 responses (47.6%) by imaging journals, but 6 were excluded as these journals do not publish articles on human research. There were 38 responses (36.5%) by nonimaging journals, but 1 was excluded for similar reasons. Fifteen of 24 imaging journals (62.5%) have official policies for both IRB approval and informed consent, and 32 of 37 (86.5%) and 29 of 37 (78.4%) nonimaging journals have such policies for IRB approval and informed consent, respectively. Eighteen imaging journals (75%) and 31 nonimaging journals (83.8%) always require authors to state in their submitted manuscripts whether IRB approval and informed consent were obtained for prospective studies; for retrospective studies, 15 imaging journals (62.5%) and 21 nonimaging journals (56.8%) always require authors to make such statements. Twenty imaging journals (83.3%) and 21 nonimaging journals (56.8%) never require submitted proof of IRB approval, and 4 imaging journals (16.7%) and 15 nonimaging journals (40.5%) at least occasionally do.

Conclusions The majority of imaging and nonimaging journals have official policies on IRB approval and informed consent, but the majority of both never require authors to submit proof of such approval and consent, trusting what the authors have said in the manuscript.

¹Department of Radiology, Winthrop-University Hospital, 259 First St, Mineola, NY 11501, USA, e-mail: dsk2928@pol.net; ²Radiology, Richmond, VA, USA

Reporting of Ethical Committee Approval and Patient Consent by Study Design in 5 General Medical Journals

Sara Schroter,^{1,2} Ros Plowman,² Andrew Hutchings,² and Adrian Gonzalez¹

Objective Authors are required to describe in their manuscripts ethical approval from an appropriate committee and how consent was obtained from participants when research involves human subjects. Previous studies have focused on concordance with these regulations within a single specialty or for clinical trials. We assess reporting of these protections for several study designs.

Design We reviewed a consecutive series of research articles published in *Annals of Internal Medicine*, *BMJ*, *JAMA*, *Lancet*, and *New England Journal of Medicine* between February and May 2003 for reporting of ethical approval and patient consent. We recorded: study design; ethical approval, naming of approving committee; type of consent; data source (directly from patient, obtained from medical records, etc), and whether study used data collected as part of a study reported elsewhere. Data were extracted by 2 researchers and agreement assessed. Differences in failure to report approval and consent by study design, journal, and vulnerable study population were evaluated using multivariable logistic regression and reported as odds ratios with 95% confidence intervals.

Results Ethical approval and consent were not mentioned in 31% and 47% of manuscripts, respectively (TABLE 15). Eighty-eight (27%) articles failed to report both approval and consent, of which 39 (44%) referred to another article. Failure to mention ethical approval or consent was significantly more likely in all study designs (except case-control and qualitative studies) compared with randomized controlled trials. Failure to mention approval was most common in *BMJ* and was significantly more likely than for the *New England Journal of Medicine*. Failure to mention consent was most common in *BMJ* and was significantly more likely than all other journals. No significant differences in approval or consent were found when comparing studies of vulnerable and nonvulnerable participants.

Conclusions Reporting of ethical approval and consent in randomized controlled trials has improved, but journals are less consistent in reporting this information for other study designs.

¹*BMJ* Editorial Office, BMA House, Tavistock Square, London, UK; e-mail: sschroter@bmj.com; ²Health Services Research Unit, London School of Hygiene & Tropical Medicine, London, UK

Table 15. Reporting of Study Approval and Consent by Study Design of All Journals Combined

	No. of Articles	Study Approval Not Mentioned*		Consent Not Mentioned†		Neither Approval Nor Consent Mentioned*†	
		No. (%)	Odds Ratio‡ (95% CI)	No. (%)	Odds Ratio‡ (95% CI)	No. (%)	Referenced Another Article, No. (%)
Study design							
Randomized controlled trial	102	7 (7)	1.00	13 (13)	1.00	5 (5)	5 (100)
Case-control	24	4 (17)	2.67 (0.69-10.3)	7 (29)	2.51 (0.76-8.28)	4 (17)	1 (25)
Cohort	96	38 (40)	8.46 (3.50-20.4)	58 (60)	13.3 (5.91-29.9)	36 (38)	21 (58)
Cross-sectional	91	44 (48)	12.3 (5.06-30.0)	54 (59)	11.8 (5.18-26.8)	42 (46)	11 (26)
Case report	25	NA	NA	23 (92)	134 (25.9-695)	NA	NA
Case series	12	NA	NA	10 (83)	40.3 (6.7-241)	NA	NA
Qualitative	8	1 (13)	1.17 (0.12-11.3)	3 (38)	1.02 (0.19-5.5)	1 (13)	1 (100)
Analysis of routine data	12	10 (83)	79.5 (13.7-462)	NA	NA	NA	NA
Journal							
BMJ	79	34 (46)	1.00	59 (79)	1.00	30 (43)	11 (37)
Lancet	127	31 (31)	0.59 (0.29-1.20)	62 (49)	0.11 (0.05-0.27)	30 (30)	9 (30)
JAMA	70	16 (23)	0.48 (0.21-1.08)	23 (34)	0.14 (0.06-0.34)	12 (18)	10 (83)
Annals of Internal Medicine	26	8 (32)	0.48 (0.17-1.40)	12 (48)	0.15 (0.05-0.46)	7 (29)	4 (57)
New England Journal of Medicine	68	15 (23)	0.33 (0.14-0.77)	12 (19)	0.03 (0.01-0.09)	9 (15)	5 (56)
Vulnerable population							
No	242	70 (32)	1.00	109 (46)	1.00	60 (28)	30 (50)
Yes	128	34 (31)	0.72 (0.40-1.28)	59 (48)	0.93 (0.51-1.68)	28 (27)	9 (32)
Total	370	104 (31)		168 (47)		88 (27)	39 (44)

Abbreviations: CI, confidence interval; NA, not applicable.

*Excludes case reports and case series for which approval is not required.

†Excludes studies reporting the analysis of routine data for which informed consent is not required.

‡From multivariable logistic regression model including design, journal, and vulnerable population.

Instructions for Authors

The Impact of Editorial Guidelines on the Classification of Race/Ethnicity in the *BMJ*

George T. H. Ellison,¹ Mikey Rosato,² and Simon Outram¹

Objective In 1996 the *BMJ* introduced “guidelines on the use of ethnic, racial, and cultural descriptions in published research.” These proposed that the categories used should “relate to the type of hypothesis under investigation” and that authors should “describe the logic behind their ‘ethnic groupings.’” The guidelines specifically recommended that authors use: “accurate descriptions rather than catch-all terms” to label the individuals or populations examined; terms that “reflect how the groups were demarcated”; and a range of information on ethnicity, race, and culture whenever it is unknown which is the “most important influence.” As such the guidelines tended to problematize race/ethnicity rather than viewing this as a potential analytical resource and might undermine its legitimate use as a marker of group identity or as a proxy for related biological and social characteristics, such as inequalities in health resulting from discrimination on the grounds of race/ethnicity.

Design To assess whether the guidelines had any impact on the classification and use of race/ethnicity in published research, the journal’s online archive was searched for any articles classified as “papers” containing the terms race or racial or ethnic* that had been published between January 1, 1994, and December 31, 2000. A random sample of 20 similar articles published between January 1, 2001, and December 31, 2004, was used to assess whether there had been any recent changes in practice. Both sets of articles were subjected to detailed content analyses to establish whether their use complied with the 1996 guidelines. Finally, the use of race/ethnicity in articles judged to address “inequalities in health” was examined in a random sample of 40 articles (classified as “papers”) published from 1994 through 2004 that contained the terms equal* or inequalit* or equit* or inequit*.

Results Between 1994 and 2000, 68 different labels were found in the 201 articles that mentioned race/ethnicity, and additional searches on these labels located another 213 articles for inclusion in the content analyses. Very few of the articles explicitly described why race/ethnicity was considered relevant to their hypothesis or analyses, and this made it difficult to assess whether appropriate categories had been used. Even after the guidelines had been introduced, few of the articles published between 1996 and 2004 explained why they had chosen to use particular categories, and none consistently used descriptive labels that indicated how these categories had been applied. There was no evidence that the guidelines influenced the prevalence of race/ethnicity in articles published between 1994 and 2004, and while the use of race/ethnicity continued to increase over this period, the proportion using race rather than ethnicity declined. Race/ethnicity did appear in articles judged to address “inequalities in health,” but it was not always mentioned in contexts for which its use would have been appropriate.

Conclusions While it is clear that the *BMJ*’s “guidelines on the use of ethnic, racial, and cultural descriptions in published research” have not been followed, it is not clear whether they were simply ignored or proved impossible to carry out. Given that the prevalence of race/ethnicity continued to increase after the guidelines

were introduced, it seems likely that the guidelines failed to deter or facilitate its use in published articles. Further research is required to assess the acceptability and feasibility of such guidelines among editorial staff, manuscript reviewers, and authors and whether guidelines should support the legitimate use of race/ethnicity as well as problematizing its inappropriate application.

¹St George’s, University of London, Grosvenor Wing—FHSS, Cranmer Terrace, London SW17 0RE, UK, e-mail: g.ellison@hscs.sgul.ac.uk; ²Social Science Research Unit, Institute of Education, University of London, London, UK

An Analysis of the Content of Medical Journals’ Instructions for Authors

David L. Schriger,^{1,2} Sanjay Arora,² Varda A. Schriger,² and Douglas G. Altman¹

Objective The CONSORT was developed and adopted with the goal of standardizing the quality of reports of randomized controlled trials. One might expect that journal instructions for authors would be similar with the goal of standardizing the quality of all medical research articles. In this study we characterize the content of instructions for authors of major medical journals.

Design We examined the content of the online versions (obtained between July and November 2004) of the instructions for authors of the top 15 general medical journals and the first- and fifth-ranked journals in 10 randomly selected specialties according to their 2001 scientific impact factor. Two abstractors independently assigned the content into 4 major and 18 subcategories and counted the total words devoted to each category. Interrater reliability of the classification was assessed.

Results Ninety-five percent (95% confidence interval, 94%–96%) of words were categorized into the same subcategory by both reviewers (TABLE 16). The 35 instructions for authors varied greatly in length (mean number of words, 3308; median, 2283; range, 885–18 927; the 4 highest word counts were 5849, 6204, 13 107, 18 927). With the exception of the lengthiest instructions, cosmetics are emphasized over content. Only 57% of instructions for authors provided guidance on the content of the article. Six of 18 subcategories were mentioned by less than half of the journals. The instructions of specialty and general journals were quite similar and the journal’s rank was not correlated with its content.

Conclusions There is little standardization of the instructions for authors among journals, and most of the attention is on cosmetic issues.

¹Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, UK; ²University of California, Los Angeles Emergency Medicine Center, University of California, Los Angeles School of Medicine, 924 Westwood Blvd, No. 300, Los Angeles, CA, 90024-2924, USA, e-mail: schriger@ucla.edu

Table 16. Presence of Categories, Mean Words per Category, and Percentage of Words per Category for 35 Journals' Instructions for Authors Stratified by Total Word Count

Category	No. (%) of Instructions Included	Mean No. of Words	All Words, %			
			All Journals (n=35)	<2000 Words (n=15)	2000-9999 (n=18)	≥10 000 (n=2)
Format	35 (100)	1733	52	67	62	24
Abstract	34 (97)	146	4	5	5	3
Reference style	32 (91)	212	6	10	7	2
How to format for submission	34 (97)	446	13	19	17	4
Logistics for how to submit	35 (100)	929	28	33	33	15
General information	35 (100)	538	16	16	15	18
About the journal	34 (97)	123	4	7	3	2
Types of articles the journal publishes	29 (83)	415	13	9	12	15
Specific information	34 (97)	686	21	15	19	29
Journal's right to access data	8 (23)	8	0	0	0	0
Journal's use of WorldWide Web	10 (29)	52	2	0	1	3
Handling the media embargo date	10 (29)	57	2	1	1	4
Duplicate publication	29 (83)	71	2	2	2	3
Reference to CONSORT guideline	28 (80)	87	3	2	3	4
Authorship definition	26 (74)	93	3	2	3	4
Peer review process	23 (66)	97	3	1	2	6
Human/animal subjects	27 (77)	97	3	2	3	2
Conflict of interest	28 (80)	123	4	2	4	5
Content of the submission	20 (57)	351	11	3	4	30
Tables and figures	10 (29)	15	0	0	0	1
Methods section	14 (40)	129	4	1	2	10
General	12 (34)	207	6	1	2	18
Total		3308	100	100	100	100

Abbreviation: CONSORT, Consolidated Standards for Reporting of Trials.

Detailed Instructions to Authors Are Required in Chinese Medical Journals

Qian Shou-chu

Objective The writing or publication frequency of detailed instructions to authors in each issue reflects the emphasis on the scientific quality of manuscripts submitted to the journals and a service for their potential contributors. This study was undertaken to investigate the neglect of the importance of instructions to authors by editors or chief editors of medical journals at the Chinese Medical Association (CMA).

Methods The instructions to authors collected from 16 journals published by the CMA were reviewed thoroughly in terms of the following well-discussed topics: financial conflict of interest, authorship or author contribution, responsibility and accountability, scientific misconduct (bias, plagiarism, fabrication), ethical issues, research integrity (design, implication of CONSORT for

randomized controlled trials and QUORUM for meta-analyses, statistical requirements, etc), grant proposals, which are frequently described in the instructions to authors of western medical journals. The 16 randomly selected journals included *National Medical Journal of China* (a general medical journal in Chinese), *Chinese Medical Journal* (a general medical journal in English), *Chinese Journal of Pathology*, *Chinese Journal of Laboratory Medicine*, *Chinese Journal of Neurology*, *Chinese Journal of Pediatrics*, *Chinese Journal of Psychiatry*, *Chinese Journal of Obstetrics and Gynecology*, *Chinese Journal of Surgery*, *Chinese Journal of Ophthalmology*, *Chinese Journal of Tuberculosis and Respiratory Diseases*, *Chinese Journal of Cardiology*, *Chinese Journal of Radiology*, *Chinese Journal of Stomatology*, *Chinese Journal of Preventive Medicine*, and *Chinese Journal of Internal Medicine*. All of these journals are core journals in China and have long been highly respected in China.

Results The instructions to authors of the 16 journals contain almost all mechanical requirements or basic elements required for the preparation of a research report or other manuscripts. But only 5 journals adequately describe the types of articles including research design (randomized controlled trial, meta-analysis, etc), performance, data collection and analysis, statistical methods, implications and conclusions in the instructions to authors. Four (*National Medical Journal of China*, *Chinese Medical Journal*, *Chinese Journal of Laboratory Medicine*, and *Chinese Journal of Neurology*) of the 5 journals specifically emphasize the requirements for conflict of interest, research ethics (report of informed consent and ethics committee approval in clinical trials), and other issues concerning research misconduct, grant proposals, authorship or author contributions, as well as responsibility and accountability.

Conclusions The published instructions to authors in the journals of the CMA are inadequately prepared or somewhat outdated. In the past years, these instructions have been less helpful in guiding potential contributors in the preparation of their research papers, thus inhibiting the improvement of the quality of the medical journals. The revision of the existing instructions should be timely, thereby, editors may specify the present editorial requirements and the answers to the queries raised by contributors or add the current results of worldwide editorial research to the instructions.

Chinese Medical Journal, 42 Dongsi Xidajie, Beijing 100710, China, e-mail: qsc@ht.rol.cn.net

Open Access

The Impact of Author Page Charges on Published Research in Infectious Diseases

Surabhi Liyanage¹ and C. Raina MacIntyre²

Objective The question of who pays for research to be conducted and published is an important one because it may influence the nature of what is and is not published. Funding to conduct research may come from the pharmaceutical industry, research institutions, government, or philanthropic bodies. The traditional model of medical publishing has relied on subscriptions for funding. There has been increasing interest in making the results of scientific research freely available. One proposed mechanism is an author-pay system of publishing, which shifts the cost from subscribers to

authors. The aims of this study were to study the impact (if any) of author page charges on the nature, funding, and type of published research and to determine the association of industry funding with types of published research.

Design Four journals of infectious diseases with comparable scope were studied, 2 with page charges and subscription (mixed model) and 2 that rely on subscriptions alone. Key variables included type of research study (clinical trial, basic science, public health), area of research, author demographics (developed vs developing countries), study setting (developed vs developing countries), and industry funding.

Results We extracted data from 463 research articles in 4 journals. Authors from developing countries were significantly less likely to be published in the mixed-model journals (odds ratio [OR], 0.25; 95% confidence interval [CI], 0.15-0.41; $P < .001$). Industry-funded studies were not significantly different between mixed-model journals and subscription-model journals. However, clinical trials published in any type of journal were significantly more likely to be industry funded than any other type of research (OR, 12.7; 95% CI, 7.0-22.9; $P < .001$). Industry-funded research was significantly less likely to be about diseases predominantly affecting those living in the developing countries (OR, 0.47; 95% CI, 0.25-0.89; $P < .05$).

Conclusions We cannot exclude an impact of an author-pay system on the nature of published research. Shifting publishing costs to authors may favor well-funded organizations, industry-sponsored research, and wealthy countries. Although we did not find that industry-funded research was more common in mixed-model journals, a relationship clearly exists between industry funding and certain types of published research. Our study only looked at journals that operate on a mixed model and have minimal charges. Entirely nonsubscription models will place an even greater financial burden on authors, thus disadvantaging authors from developing countries or from poorly funded organizations. This needs to be considered when planning for open-access models.

¹University of Sydney, Faculty of Medicine, 2/21-23 Queens Rd, Westmead, Sydney, NSW, 2145, Australia, e-mail: siliyanag@gmp.usyd.edu.au; ²National Centre for Immunisation Research and Surveillance of Vaccine Preventable Diseases, Children's Hospital at Westmead, Westmead, Sydney, Australia

Peer Review Process

Do You Need to Be an Editor to Accept or Reject Research Papers Sent to the *BMJ*?

Gary Bryan, John Fletcher, and Rajendra Kale

Objective To find out if a nonqualified but experienced member of the *BMJ* editorial administrative team (G.B.) can make correct decisions on research papers submitted to the *BMJ*.

Design Prospective observational study comparing the accuracy of the decisions made by G.B. against the gold standard of the decision made by *BMJ* editors. Sample size: we wished to estimate the ability of the staff member correctly to predict papers that are accepted with a 95% confidence interval (CI) of plus or minus

10%. Assuming the *BMJ* acceptance rate would remain at 7% and that the staff member would identify 50% of accepted papers, we calculated that we would need a sample size of 1372 papers.

Results At the beginning of January 2005, G.B. had judged 1496 papers. *BMJ* editorial had accepted 84 papers, rejected 1377, and had yet to make a final decision on 35. G.B. decided to accept 37 of the accepted papers (sensitivity 44%, 95% CI, 33%-55%) and decided to reject 1237 of the rejected papers (specificity 90%, 95% CI, 88%-91%).

Conclusions If *BMJ* replaced its editorial team with a single staff member to make decisions on papers it would reject more than half of the papers it currently considers publishable; it would also accept 10% of the papers it currently rejects.

BMJ Editorial Office, BMA House, Tavistock Square, London WC1H 9JR, UK, e-mail: gbryan@bmj.com

Qualitative Profile of Journal Peer Reviewers and Predictors of Peer Reviewer Quality

Michael Callaham and John Tercier

Objective To perform an in-depth qualitative assessment of the attitudes and value systems of journal peer reviewers.

Design Potential participants were reviewers at *Annals of Emergency Medicine* who had completed at least 5 reviews in the past 2 years. After pilot testing, a blinded randomized stratified sample of 72 was enrolled to undergo structured recorded interviews. The data collected was analyzed using grounded theory (Glaser, Strauss, and Clark), to produce a qualitative profile of their attitudes toward the peer review process. Categorized responses included the benefits to reviewers of peer review, how their reviewing has changed with experience or past training, how they structure their review process, how they assess the most useful crossover skills, and others.

Results The reviewers had an average of 9 years of experience reviewing; 90% also reviewed for other journals, 63% were on editorial boards, and 57% had performed National Institutes of Health or equivalent grant peer review. Their quality scores were typical of the total reviewer pool and all academic ranks were included. Many reviewers challenged conventional beliefs about the purpose and process of peer review, and showed concern over tensions within the process; only a few issues are listed herein: purpose of review, evaluation of the manuscript for the journal was seen as being in tension with education of the researcher, with the latter taking precedence for the majority; motivation. "Duty" was constructed as both an ethical and a systems imperative, based on a strong empathetic identification by the reviewer with the author; skills acquisition, formal training was downplayed in preference to editorial feedback, which was believed to be insufficient and also perceived as a form of direct reward; grants review and institutional review board membership were considered the most useful form of related activity; validity, peer review's main strength was seen as being based less on the skills of the individual reviewers and more on its corporate nature as a system of checks and balances; and defects, although there was concern over "poor" science being frequently published, a larger concern was that of "good" science being "strangled in its cradle" by the process.

Conclusions Structured interviews demonstrated that the motivation and desires of peer reviewers differ from those of editors and should help provide insights for improving the process.

Department of Anthropology, History and Social Medicine and Division of Emergency Medicine, University of California, San Francisco, Box 0208, San Francisco, CA 94143-0208, USA, e-mail: mlc@medicine.ucsf.edu

A Randomized Trial on the Effect of Statistical Reviewing and Checklist on Manuscript Quality

Erik Cobo, Albert Selva, Josep Maria Ribera, Francesc Cardellach, Ruth Dominguez, Agustín Urrutia, Vicens Fonollosa, Mercedes Belmonte, Celestino Rey-Joly, and Miquel Vilardell

Objective To estimate the effect of adding statistical experts and checklists, such as CONSORT or STARD, to the peer review process.

Design A total of 115 original research papers consecutively selected between May 2004 and March 2005 for peer review by the *Medicina Clínica* editorial committee. This study used a factorial 2 × 2 with no statistical expert and no checklist as controls. Manuscripts were randomized to add 1 statistical reviewer to the clinical peer review process and to provide checklists to reviewers. (Sample size rationale: 100 papers provide an 80% power to test a 55% standardized difference.) The outcome measured was quality of paper, from draft to final version, according to Goodman Scale (sum over all specific items) assessed by 2 blinded evaluators.

Results Of the 115 original research papers sent to reviewers, 16 (13.9%) were rejected. Twenty-one (18.3%) were about interventions, 46 (40.0%) were longitudinal designs, 28 (24.3%) cross-sectional, and 20 (17.4%) others. Rejected papers had a lower initial mean (SD) quality on the overall Goodman Scale (3.75 [2.22]) than the accepted papers (4.49 [1.78]). The estimated effect of adding 1 statistical reviewer was 5.49 (95% confidence interval [CI], 3.05–7.94) and the effect of sending a checklist 0.87 (95% CI, –1.57 to 3.32) with no interaction between them (1.11 [95% CI, –1.33 to 3.56]).

Conclusions This study is the first blinded prospective randomized study about the effect of adding both statistical reviewers and guidelines. It raised the positive effect of the statistician on a direct measure of the paper quality but failed to establish a positive effect of the guidelines. This lack of such effect can be explained by the possible use of the guidelines on the control group.

Medicina Clínica, Doyma, EIO, FME, UPC, C/ Pau Gargallo 5, Barcelona 08028, Spain, e-mail: erik.cobo@upc.edu

Blinded vs Unblinded Peer Review in a Non-English-Language Journal: A Randomized Controlled Trial

Torben V. Schroeder,^{1,2} Ole H. Nielsen,¹ for the Editorial Team of *Ugeskrift for Læger* (Journal of the Danish Medical Association)

Objective To examine the effect on quality of peer review in a national non-English-language journal when the identity of reviewers is revealed to the authors.

Design Consecutive eligible papers were forwarded to 2 reviewers randomized to either have their identity revealed to the authors or to remain anonymous. Using a validated instrument consisting of 8 items, each scored on a 5-point Likert scale, 2 editors rated the quality of reviews independently and in a blinded fashion. Based on power calculations (MERIDIF, 0.3; power, 0.9; and $\alpha=.05$), the goal was to use 161 manuscripts, which demanded 190 manuscripts with a 20% loss. Additionally, a questionnaire survey with reviewers of all enrolled manuscripts was undertaken to assess their views of an open peer review.

Results By May 2005, 190 manuscripts had been included, and 182 completed data sets were available for analysis. There was no significant difference in total mean scores between the quality, as assessed by the editors of reviews from identified vs anonymous reviewers (3.34 vs. 3.28, respectively). However, the reviews produced by identified reviewers were judged to be superior for all 8 items ($P<.01$) when they were considered together. However, none reached individually the level of significance (TABLE 17). One third of the reviewers indicated in their response to the questionnaire that they felt more comfortable when their identity was unknown to the author(s); and significantly more among reviewers randomized to anonymous review compared with open review (43% vs 25%; $P<.001$). Only 9%—equal in both groups—indicated that open review would make them modify their review report, in most instances to be more meticulous.

Table 17. Effects of Reviewers Being Randomly Assigned to be Identified on the Quality of Their Review: Means of 2 Editors' Assessment

Item*	Mean (SD)		Difference (SD)
	Identified Reviewers (n=182)	Anonymous Reviewers (n=182)	
1. Importance	3.32 (0.87)	3.30 (0.79)	0.01 (1.19)
2. Originality	3.16 (0.83)	3.12 (0.84)	0.04 (1.19)
3. Method	3.39 (0.91)	3.36 (0.90)	0.03 (1.33)
4. Presentation	3.48 (0.96)	3.36 (0.91)	0.13 (1.30)
5. Constructiveness of comments	3.53 (0.90)	3.45 (0.85)	0.08 (1.31)
6. Substantiation of comments	3.27 (1.00)	3.20 (0.94)	0.07 (1.37)
7. Interpretation of results	3.13 (0.89)	3.12 (0.84)	0.01 (1.19)
8. Tone of review	3.44 (0.53)	3.32 (0.60)	0.12 (0.86)
Mean total score of items 1-7	3.34 (0.82)	3.28 (0.78)	0.06 (1.22)

*Items scored on a 5-point scale (1, poor; 5, excellent) from van Rooyen S, Godlee F, Evans S, Black N, Smith R. Effect of open peer review on quality of reviews and on reviewers' recommendations; a randomised trial. *BMJ*. 1999;318:23-27.

Conclusions In contradiction to results of existing international trials on the effect of open review on quality, this study indicates a minor although significant advantage of an open peer review process. However, this observation should be balanced against the finding that one third of reviewers felt more comfortable being anonymous.

¹*Ugeskrift for Læger (Journal of the Danish Medical Association)*, Copenhagen, Denmark; ²Dept Vascular Surgery RK, Rigshospitalet 3111, Blegdamsvej 9, DK 2100 Copenhagen Ø, Denmark, e-mail: torben.schroeder@rh.hosp.dk

Consistency Between Reviewers and Editors About Which Papers Should Be Published

James R. Scott,^{1,2} Sheryl Martin,¹ and Leon Burmeister³

Objective Prepublication peer review of manuscripts submitted to medical journals is a crucial part of the scientific process, but questions remain about what constitutes a good review and how reviews affect the ultimate fate of the paper. The purpose of this study was to evaluate interreviewer reliability and the effect it has on acceptance or rejection of manuscripts.

Design Original research reports submitted to *Obstetrics & Gynecology* from January 1, 2003, through December 31, 2004, were included if they were evaluated by at least 2 expert reviewers, an editorial board member, and 1 of the 3 editors. Other manuscripts, such as systematic reviews, commentaries, case reports, and those evaluated by fewer than 3 reviewers, were excluded. The major statistical analysis was based on estimation of reliability among expert reviewers, editorial board reviewers, and the 3 editors. Reliability coefficients were estimated for continuous measures of the overall grading on a 0-100 scale. Categorical results were evaluated by κ statistics to quantitate agreement based on recommendations classified as acceptance, minor revision, major revision, and rejection.

Results Of 964 manuscripts analyzed, agreement among expert reviewers was 37.1% and the κ value among all 3 reviewers was 0.293. The correlation coefficient for manuscript grades assigned by 3 reviewers was 0.366. Agreement between editor decisions and the expert reviewer assessment was 47.4% ($n=484$), 47.3% ($n=245$), and 45.5% ($n=235$), and was 47.0% when the results were pooled. Agreement between editors and editorial board reviewers was 57.6%, 58.6%, and 61.5%, respectively, and 58.8% when results were pooled.

Conclusions Overall, agreement among peer reviewers on which individual research papers should be published was relatively low. Interreviewer consistency was lowest among expert reviewers and highest between editors and editorial board members and when all reviewers agreed on the rating.

¹*Obstetrics & Gynecology*, Washington, DC, USA; ²Department of Obstetrics and Gynecology, University of Utah School of Medicine, 423 Wakara Way, Suite 201, Salt Lake City, UT 84108-1242, USA, e-mail: jscott@hsc.utah.edu; ³College of Public Health, E220G General Hospital, Iowa City, IA, USA

Are Reviewers Suggested by Authors as Good as Those Chosen by Editors? Results of a Rater-Blinded, Retrospective Study

Elizabeth Wager,¹ Emma C. Parkin,² and Pritpal S. Tamber²

Objective BioMed Central requires authors to suggest 4 reviewers when making a submission. The review process is open—authors and reviewers know each other's identity—although reviewers can make confidential comments to editors. Reviews are published alongside accepted articles so readers may know the reviewers' identity and their recommendations. The objective of this study was to compare the performance of reviewers suggested by authors with those chosen by editors in terms of review quality and recommendations about submissions in an online-only medical journal.

Design Pairs of reviews from 100 consecutive submissions to BioMed Central (from November 2003 to April 2004, with 1 author-nominated and 1 editor-chosen reviewer and a final decision) were scored by 2 raters, blinded to reviewer type, using a validated review quality instrument that rates 7 items on 5-point Likert scales. The raters discussed their ratings after the first 20 pairs (keeping reviewer type masked) and resolved major discrepancies in scoring and interpretation to improve interrater reliability. Reviewers' recommendations were also compared.

Results Reviewer source had no impact on review quality (mean score, 2.24 [SD, 0.55] for reviewers suggested by authors and 2.34 [SD, 0.54] for reviewers chosen by editors) or tone (mean scores on additional question, 2.72 for reviewers suggested by authors vs 2.82 reviewers chosen by editors; maximum score = 5 in both cases). However, author-nominated reviewers were significantly more likely to recommend acceptance (47 vs 35) and less likely to recommend rejection (10 vs 23) than editor-chosen reviewers after initial review ($P < .001$). Recommendations made by reviewers suggested by authors and those made by reviewers suggested by editors were similar (65 vs 66 recommended acceptance and 10 vs 14 recommended rejection; $P = .47$). The number of reviewers choosing "unable to decide" on acceptance/rejection was similar in both groups at both review stages.

Conclusions Author-nominated reviewers produced reviews of similar quality to those of editor-chosen reviewers but were more likely to recommend acceptance.

¹Sideview, Princes Risborough, HP27 9DE, UK, e-mail: liz@sideview.demon.co.uk; ²BioMed Central, London, UK

Publication Bias

How Does Prior Publication Affect Full Publication of Completed Clinical Trials?

Yuan-I Min,¹ Aynur Unalp-Arida,² Roberta Scherer,³ and Kay Dickersin⁴

Objective To determine the effect of prior publication during the course of a clinical trial on the rate of full publication.

Design An existing data set from a retrospective follow-up study (1988-1989) of 3 cohorts of initiated studies, evaluating publication bias, was used for this analysis. Only trials that had completed

study enrollment at the time of follow-up were included (n=306). Publication history and characteristics of the trials were based on investigators interviews. Full publication was defined as the first publication, with a length of 3 or more pages, after completion of enrollment. Prior publication was defined as any publication (eg, abstracts, design reports) prior to the full publication. Analysis of time to full publication since enrollment completion used Cox regression. Two variables represented prior publication: (1) any publication prior to completion of study enrollment (baseline) and (2) any publication after completion of study enrollment and prior to full publication (follow-up). Prior publication during follow-up was a time-dependent variable.

Results Seventy-two percent of the trials reported a full publication at the time of the interview. The median time to full publication after enrollment completion was 3 years. Prior publication was reported in 39% of the trials: 28% at baseline and 16% during follow-up. After adjusting for statistical significance of the primary results and funding sources (both were significantly associated with full publication), prior publication during follow-up was significantly associated with full publication (hazard ratio [HR], 1.85; 95% confidence interval [CI], 1.29-2.64) but prior publication at baseline was not (HR, 1.03; 95% CI, 0.76-1.40).

Conclusions In clinical trials, prior publication does not delay full publication. Prior publication was associated with a higher rate of full publication (significantly so for prior publication during follow-up), regardless of the statistical significance of primary results or funding sources.

¹Department of Epidemiology and Statistics, MedStar Research Institute, 6495 New Hampshire Ave, Suite 201, Hyattsville, MD 20783, USA, e-mail: nancy.min@medstar.net; ²Center for Clinical Trials, Johns Hopkins University, Baltimore, MD, USA; ³University of Maryland, Department of Epidemiology and Preventive Medicine, Baltimore, MD, USA; ⁴Brown University, Department of Community Health, Providence, RI, USA

Quality of Reporting Trials and Other Studies

What's in a NAME? Clinical Trial Acronyms and Research Impact

Matthew B. Stanbrook^{1,2,3} and Donald A. Redelmeier^{1,2}

Objective The use of acronyms to name clinical trials is a popular yet controversial practice among researchers. We evaluated whether the use of an acronym to name a clinical trial is associated with increased research impact.

Design Multiple studies published between 1953 and 2003, addressing each of 13 research questions, were identified from all systematic reviews completed by the Cochrane Heart Group as of January 2004. Studies were classified as having or not having an acronym name based on examination of the original publications. The impact of acronym and nonacronym studies, as measured by the article citation rate from the time of publication (Web of Science, Thomson ISI) was compared. To control for article content, studies were analyzed as clusters according to the research question they addressed, using hierarchical linear modeling.

Results Of 173 articles identified, each representing a single study, 59 (24%) had an acronym name. In a multivariable model, an acronym

name was independently associated with an increased citation rate (rate ratio, 1.69; 95% confidence interval [CI], 1.03-2.78; $P = .04$), as were adequacy of allocation concealment, researcher nationality, industrial sponsorship, and positive study outcome. In a subset of articles matched for study question and publishing journal, an acronym name remained associated with increased citations (rate ratio, 2.16; 95% CI, 1.11-4.18; $P = .02$).

Conclusions Studies with acronym names have twice the citation rate of similar studies without acronym names, independent of study size, quality, and outcome. This observation is consistent with theories in cognitive psychology regarding word perception.

¹University of Toronto, Toronto, Ontario, Canada; ²Institute for Clinical Evaluative Sciences in Ontario, Toronto, Ontario, Canada; ³Toronto Western Hospital, Seventh Floor, East Wing, 399 Bathurst St, Toronto, Ontario, Canada M4N 3M5, e-mail: m.stanbrook@utoronto.ca

Quality of Placebo-Controlled Trials of Alternative and Conventional Medicine: Matched-Pair Study

Aijing Shang,¹ Karin Huwiler,¹ Linda Nartey¹ Peter Jüni,¹ and Matthias Egger^{1,2}

Objective To compare the reported quality of placebo-controlled trials of complementary/alternative medicine (CAM) with comparable trials of conventional medicine.

Design Placebo-controlled trials of CAM (homoeopathy, Chinese herbal medicine [CHM], and Western herbal medicine [WHM]) were identified in literature searches (19 electronic databases, reference lists of relevant articles, and contacts with experts for homoeopathy and WHM; 11 databases and hand searches of 48 Chinese-language journals for CHM). We included all eligible trials of homoeopathy and CHM and a sample of eligible trials of WHM. Conventional medicine trials matched to CAM trials for condition and type of outcome were randomly selected from the Cochrane Controlled Trials Register. Assessment of study quality focused on randomization (generation and concealment of allocation sequence), blinding (patients, therapists, and outcome assessors) and analysis (intention-to-treat principle or other). Accepted definitions were used for classifying methods as adequate or inadequate. Analysis according to the intention-to-treat principle was assumed if the reported number of participants randomized and the number analyzed were identical. Trials described as double-blind with adequate randomization were classified as having higher methodological quality.

Results We included 335 CAM trials (110 of homoeopathy, 136 of CHM, and 89 of WHM) and 335 conventional medicine trials. Most trials were small (<100 patients) and assessed "soft" outcomes (for example, global treatment response or symptoms). English-language trials dominated all groups except for CHM (88% Chinese). The quality of reporting was low both for CAM and for conventional medicine (TABLE 18), indicating that methodological quality remains unclear for many trials.

Conclusions Our findings do not confirm the widely held belief that the evidence on the effectiveness of CAM is inferior to the evidence that is available for conventional medicine. Except for trials of CHM, the quality of trials in CAM is not lower than in conven-

tional medicine—quality is unacceptably low in all groups.

Table 18. Reported Quality of Placebo-Controlled Trials of Complementary/Alternative Medicine and Matched Trials of Conventional Medicine

Quality Measure	Complementary/Alternative Medicine, No. (%) (n=335)			Conventional Medicine, No. (%) (n=335)
	Homoeopathy (n=110)	Chinese Herbal Medicine (n=136)	Western Herbal Medicine (n=89)	
Concealment of allocation				
Unclear	46 (42)	123 (90)	55 (62)	264 (79)
Inadequate	15 (13)	1 (1)	9 (10)	24 (7)
Adequate	49 (45)	12 (9)	25 (28)	47 (14)
Generation of allocation sequence				
Unclear	65 (59)	80 (59)	48 (54)	246 (73)
Inadequate	18 (16)	34 (25)	8 (9)	16 (5)
Adequate	27 (25)	22 (16)	33 (37)	73 (22)
Blinding described as*				
Double-blind	101 (92)	56 (41)	83 (93)	309 (92)
Assessor-blind	14 (13)	20 (15)	18 (20)	64 (19)
Analyzed according to intention to treat	33 (30)	85 (63)	24 (27)	101 (30)
High quality	21 (19)	3 (2)	18 (20)	23 (7)

*Categories for blinding are not exclusive.

¹Department of Social and Preventive Medicine, University of Berne, Finkenhubelweg 11, CH-3012 Berne, Switzerland, e-mail: egger@ispm.unibe.ch; ²Department of Social Medicine, University of Bristol, Bristol, UK

Improving the Quality of Reporting of Randomized Controlled Trials Evaluating Herbal Interventions: Implementing the CONSORT Statement

Joel J. Gagnier,¹ Heather Boon,² Paula Rochon,³ David Moher,⁴ Joanne Barnes,⁵ and Claire Bombardier⁶

Objective Herbal medicinal products are widely used, vary greatly in content and quality, and are actively tested in randomized controlled trials (RCTs). Therefore, RCTs testing herbal medicine interventions must clearly report the specifics of the intervention. Our objective was to develop recommendations for reporting RCTs of herbal medicine interventions.

Design We identified and invited potential participants with expertise in clinical trial methods, clinical trial reporting, pharmacognosy, herbal medicinal products, medical statistics, and herbal product manufacturing to participate in telephone discussion and a consensus meeting. Three phases were conducted: (1) premeeting item generation via telephone; (2) consensus meeting; and (3) postmeeting feedback. Sixteen experts participated in premeeting telephone calls for item generation and 14 participants attended a consensus meeting in Toronto, Ontario, in June 2004. During the consensus meeting, a modified Delphi technique was used to aid discussion and debate of information required for reporting RCTs

of herbal medicines.

Results After extensive discussion, the group decided that context-specific elaborations of 9 CONSORT items to RCTs of herbal medicines were necessary: items 1 (title and abstract), 2 (background), 3 (participants), 4 (interventions), 6 (outcomes), 15 (baseline data), 20 (interpretation), 21 (generalizability), and 22 (overall evidence).

Conclusions The elaboration of item 4 of the CONSORT statement outlines specific information required for complete reporting of the herbal medicine intervention. The reporting suggestions presented will support clinical trialists, editors, and reviewers in reporting and reviewing RCTs of herbal medicines and will help readers in interpreting the results.

¹Department of Health Policy, Management and Evaluation, Department of Medicine, University of Toronto, 5955 Ontario St, Unit 307, Windsor, Ontario, Canada N8S1W6, e-mail: j.gagnier@utoronto.ca; ²Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada; ³Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada; ⁴Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; ⁵Centre for Pharmacognosy and Phytotherapy, School of Pharmacy, University of London, London, UK; ⁶Institute for Work and Health, Toronto, Ontario, Canada

An Analysis of the Quality of Reporting of General Surgical Randomized Controlled Trials Published in General Health Care and Surgical Journals

Martin Wiener,¹ Sabapathy Prakash Balasubramanian,¹ Zeiad Kaid,¹ Ravindranath Tiruvoipati,² Diana Elbourne,³ and Malcolm Walter Reed¹

Objective To compare the quality of reporting of general surgical randomized controlled trials (RCTs) published in high-profile general health care and surgical journals, with particular attention to adherence to the CONSORT guidelines, and to identify factors influencing quality.

Design The study comprised an observational, cross-sectional evaluation of general surgical RCTs published in 10 high-profile, English-language general health care and surgical journals in 2003. All manuscripts appearing in these journals and relating to RCTs on general surgical topics were selected (82 manuscripts). Thirteen manuscripts were excluded, mainly because they related to only part of an RCT or because the trial was not truly randomized. For each RCT, the characteristics were noted, then masking was carried out for authorship, institutions, statistician involvement, industry funding, and journal of publication. Jadad score and allocation concealment were recorded for each RCT. Two observers then independently assessed the masked RCTs for quality of reporting using a 30-item checklist based on current CONSORT guidelines.

Results The results of quality assessment by Jadad score, allocation concealment, and CONSORT checklist are shown in TABLE 19. Univariate analysis demonstrated that RCTs with higher author numbers ($P = .03$), multicenter studies ($P = .002$), and studies with a declared funding source ($P = .02$) were of significantly better quality.

Table 19. Results of the Quality Assessment

Assessment Criteria	General Health Care Journals (n=8)	Surgical Journals (n=61)	P Value
Jadad score, No. (%)			
Low (<3)	2 (25)	41 (67)	.046*
High (≥3)	6 (75)	20 (33)	
Allocation concealment, as assessed by Schulz et al, No. (%)			
Unclear/inadequate	4 (50)	50 (82)	.06*
Adequate	4 (50)	11 (18)	
Modified CONSORT score, median (interquartile range)	85.45 (81.09-86.13)	68.97 (62.89-73.11)	<.001†

*By Fisher exact test.
†By Mann-Whitney U test.

Conclusions The quality of reporting of general surgical RCTs published in high-profile general health care journals and in journals endorsing the CONSORT statement is significantly higher than for RCTs published in high-profile surgical journals that do not endorse the CONSORT statement. More widespread endorsement of the CONSORT statement may lead to an increase in the quality of reporting of RCTs.

¹Academic Unit of Surgical Oncology, University of Sheffield, 54 Blair Athol Rd, Sheffield S11 7GB, UK, e-mail: martinwiener@fsmail.net; ²ECMO Unit, University of Leicester, Leicester, UK; ³Department of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, University of London, London, UK

Quality of Journal Articles

Titles of Articles in Peer-Reviewed Journals Lack Essential Information: A Structured Review of Contributions to 4 Leading Medical Journals, 1995 and 2001

Paul Z. Siegel,¹ Stephen B. Thacker,² Richard A. Goodman,³ and Cathleen Gillespie¹

Objective Information about study methods can help readers judge at a glance their level of interest in a given article. We assessed the information content of titles from articles in 4 leading medical journals according to a specified typology—Methods, Results, Conclusion, Specified Data Set, or Topic Only—and measured change in the information content between 1995 and 2001.

Design Three medical epidemiologists applied the above typology to the titles of 417 original research articles published in the *BMJ*, *JAMA*, *Lancet*, and *New England Journal of Medicine* from July through December 2001. The authors determined the percentage of titles that fit each of the typology categories. Using the χ^2 test of independence, we compared these percentages with the corresponding percentages among titles from 420 original research articles published in the same journals from July through December 1995.

Results Among titles of articles published during July through December 2001, information about study methods appeared most frequently in the *BMJ* (96%) vs 20% to 41% among the other journals (TABLE 20). The percentage of titles in 2001 was higher than in 1995 only at the *BMJ* (96% vs 49%; $P < .001$).

Conclusions The substantial increase from 1995 to 2001 in the percentage of *BMJ* titles that contain information about study methods is consistent with that journal's stated editorial policy. Since November 2003 the International Committee of Medical Journal Editors (ICMJE) has recommended that titles of research articles include information about study design. Consensus criteria on what constitutes key information about study design should help additional journals implement the ICMJE recommendation.

¹National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Hwy, Mail Stop K-30, Atlanta, GA 30341, USA, e-mail: pzs1@cdc.gov; ²Office of Workforce and Career Development, ³Office of the Chief of Public Health Practice, Centers for Disease Control and Prevention, Atlanta, GA, USA

References and Citations

Editors' Impact on Improving the Accuracy of References: Randomized Comparison of Standard Practice, Brief Reminder, and Instructional Intervention

Kristina Fišter,^{1,2} Ana Marušić,³ Andrew Hutchings,⁴ Josipa Kern,¹ and Matko Marušić³

Objective To investigate whether editor intervention prompts authors to reduce the number of errors in references of manuscripts they submit for publication.

Design We randomly allocated 75 consecutive manuscripts accepted for publication in a general medical journal to 1 of 3 interventions.

Table 20. Distribution of Titles, by Category, in Articles Published During July-December 1995 and 2001*

	BMJ			JAMA			Lancet			N Engl J Med			Total		
	1995,% (n= 133)	2001,% (n= 112)	P Value	1995,% (n= 73)	2001,% (n= 95)	P Value	1995,% (n= 111)	2001,% (n= 118)	P Value	1995,% (n= 103)	2001,% (n= 92)	P Value	1995,% (n= 420)	2001,% (n= 417)	P Value
Topic Only	35	4	<.001	37	49	.02	41	64	.005	48	76	.05	40	47	.13
Methods	49	96	.001	38	42	.15	34	26	.40	31	20	.05	39	47	.07
Results	23	0	<.001†	23	4	.005	25	3	<.001	22	4	<.001	24	3	<.001
Conclusion	0	0	‡	1	1	‡	1	0	‡	0	0	‡	0	0	‡
Data Set	2	2	‡	8	9	.44	5	11	.11	0	0	‡	4	6	.15

*Sum of categories is greater than 100% for most columns because some article titles contain information about more than 1 element (eg, Methods and Results).

†Cell counts increased by a factor of 1 to allow for χ^2 test.

‡ χ^2 Test could not be performed due to inadequate expected cell counts (< 5).

The editor-in-chief returned manuscripts to the authors for final changes with a cover letter corresponding to the intervention group: standard practice (prompting authors to acknowledge required changes, with no specific mention of the references section), brief reminder (standard practice plus a sentence prompting authors to pay special attention to the accuracy of references), or instructional intervention (standard practice plus a paragraph highlighting the importance of the accuracy of references and a copy of reference citation formats recommended by the International Committee of Medical Journal Editors [ICMJE]). In a manner that was blinded as to intervention group, errors were classified as technical (eg, punctuation) or substantive (eg, misspelled author's last name) compared with the manuscript editor's standard. Differences in the accuracy and rate of errors in revised manuscripts between standard practice and the other interventions were examined using logistic and Poisson regression. We adjusted for ICMJE reference format and errors in the original manuscript. The manuscripts were the unit of analysis and included as random effects.

Results We considered 2035 pairs of references after excluding 12 references deleted from original manuscripts and 79 added to revised manuscripts. The percentages of completely accurate references and references without technical and substantive errors in the original manuscripts were 14.9%, 30.1%, and 42.7%, respectively (TABLE 21). Small but statistically significant improvements in completely accurate and technically correct references were observed in the instructional intervention group compared with standard practice. These improvements were also better than in the brief reminder group ($P = .02$ and $P < .01$, respectively). No statistically significant differences were observed for improvement in substantive errors.

Table 21. Comparison of Reference Quality Before and After Interventions

Quality Measures	Standard Practice (25 Manuscripts; n=720 References)		Brief Reminder (25 Manuscripts; n=613 References)		Instructional Intervention (25 Manuscripts; n=702 References)	
	Before	After	Before	After	Before	After
Completely accurate references, No. (%)	109 (15.1)	105 (14.6)	94 (15.3)	95 (15.5)	100 (14.2)	139 (19.8)
References with no technical errors, No. (%)	257 (35.7)	251 (34.9)	161 (26.3)	162 (26.4)	195 (27.8)	240 (34.2)
References with no substantive errors, No. (%)	274 (38.1)	283 (39.3)	286 (46.7)	302 (49.3)	308 (43.9)	337 (48.0)
Substantive errors, total No. (mean per reference)	743 (1.03)	715 (0.99)	484 (0.79)	445 (0.73)	643 (0.92)	583 (0.83)

Conclusions Editors' instructional intervention can produce small improvements in the accuracy of references as supplied by authors.

¹Andrija Stampar School of Public Health, Zagreb, Croatia; ²BMJ, BMA House, Tavistock Square, London WC1H 9JR, UK, e-mail: kfister@bmj.com; ³Croatian Medical Journal, Zagreb University School of Medicine, Zagreb, Croatia; ⁴Health Services Research Unit, London School of Hygiene and Tropical Medicine, London, UK

Web Citations: Going, Going, Still There

Gunther Eysenbach, M. J. Suhonos, and Jean-Sebastian Dumais

Objective Citations of Web pages in medical and scientific publications are becoming more and more common; however, Web sites are transient and can become unavailable overnight. The problem of unstable Web citations has recently been referred to as an issue "calling for an immediate response" by publishers and authors. Services such as the Internet Archive and Google offer archiving (caching) of Internet documents, but this is done randomly, does not focus on academic references, and cannot be initiated by authors, editors, or publishers wanting to cache a specific Web reference.

Design We developed and pilot-tested a novel tool called Webcite (<http://www.webcitation.org>) designed to be used by authors, readers, editors, and publishers, allowing them to permanently archive and retrieve cited Internet references. The process can be initiated by the author (or publisher) of a citing manuscript, who can upload a manuscript to the Webcite server. This initiates the Webcite tool to automatically archive cited URLs, associated with a time stamp. The Webcite software also modifies all URL citations in the manuscript, redirecting readers to the permanently archived cached document. Participating journal editors would ask authors to cache all cited URLs prospectively before submission. Webcite also works as a focused crawler, automatically discovering references and caching cited URLs retrospectively on domains hosting academic journals, which does not require that authors cache cited references before submission. We evaluated the latter functionality on BioMed Central.

Results Webcite analyzed 280 752 references from 8381 articles published in all BioMed Central journals from August 1997 to April 5, 2005. A total of 6627 (2.4%) of these references were pure URL citations (eg, not a URL of a journal article), of which 4919 were unique. Fifteen hundred seventy-one cited an entire domain (ie, a Web site as opposed to a specific Web page); 2938 cited an HTML page, 222 cited a PDF file; and 15 cited .txt/.doc-extension files. Obeying a variety of robot-exclusion standards and "no-archive"/"no-cache" metatags or copyright restrictions, we succeeded in archiving 3198 (65%) of 4919 Web pages. Five hundred were not cached because of robot exclusions, but only 8 had no-archive and 7 had no-cache restrictions. Only 57 had machine-readable copyright notices. The remaining Web pages could not be cached because they were already inaccessible or had disappeared.

Conclusions Retrospective caching has limitations; by the time references are being archived, they may have disappeared already. Prospective archiving of cited references by authors or publishers at the time the manuscript is written or published is recommended to solve the problem of unstable and changing Web citations.

Journal of Medical Internet Research, Centre for Global eHealth Innovation, Toronto General Hospital, R. Fraser Elliott Building, Fourth Floor, Room 4S435, 190 Elizabeth St, Toronto, Ontario, Canada M5G 2C4, e-mail: geysenba@uhnres.utoronto.ca

Rhetoric and Interpretation

The Rhetoric of Efficacy in Donepezil Trials

John R. Gilstad¹ and Thomas E. Finucane²

Objective Clinical interpretation of biomedical literature is modulated by the rhetoric with which experimental data are presented. The rhetoric of most reports on donepezil for Alzheimer dementia (AD) suggests significant clinical efficacy, while that of 2 others, including the latest and largest trial, suggests a distinctly more limited efficacy. The purpose of our study was to examine whether differences in rhetoric are based on differences in experimental findings.

Design From each randomized trial of donepezil for AD, we tabulated primary and secondary end-point treatment effects. We excerpted key interpretive sentences from 5 prominent textual locations in each article, defined by typical reading patterns. We compared interpretative rhetoric in these sentences with the data to which they refer, and characterized forms and patterns.

Results Numerical treatment effects were similar across the 12 articles; statistically significant treatment effects were found consistently in measures of cognition, but less so in other domains of functioning. Rhetorical tone of key sentences lies along a broad spectrum from skeptical to enthusiastic. For example, in articles reporting a similar primary outcome difference of 2 to 3 points on a 70-point scale, skeptical rhetoric portrayed the effect as small or modest (eg, "... our results demonstrate a small beneficial effect of donepezil therapy on cognitive function") Enthusiastic rhetoric portrayed the same size effect as distinctly positive (eg, "... this multinational study demonstrates that donepezil therapy is an effective and well-tolerated symptomatic treatment") The question of alternative treatments for AD was particularly polarized, with 2 skeptical reports emphasizing the need for better treatments while 1 enthusiastic article suggested that further placebo-controlled trials would "raise ethical and practical concerns." Articles from both ends of the spectrum cite the same prior literature as supporting their interpretation.

Conclusions Experimental findings in the donepezil literature are consistent, but rhetoric varies greatly. Clinical interpretation may be affected.

¹Department of Internal Medicine, National Naval Medical Center, Bethesda, MD, USA; ²Division of Geriatric Medicine and Gerontology, Johns Hopkins Bayview Medical Center, John R. Burton Pavilion, 5505 Hopkins Bayview Circle, Baltimore, MD 21224, USA, e-mail: tfinucan@jhmi.edu

Do the Conclusions Look as Good as They Seem? A Review of Quality Improvement Intervention Studies

Linda Li,¹ Lorenzo Moja,² Alberto Romero,³
and Jeremy Grimshaw¹

Objective To assess the appropriateness of conclusions reported in quality improvement (QI) intervention studies.

Design Eleven medical journals or health services research journals were hand-searched for QI intervention studies published between January 2002 and December 2003. A 38-member clinical epidemiology panel rated quotes on a 7-point Likert scale (a higher score means that authors inferred stronger causality), assuming that all quotes were from well-designed randomized controlled trials (RCTs). Two-way analysis of variance was used to assess main effects and the interaction between study designs (RCTs vs non-RCTs) and results of primary outcomes (statistically significant and mixed results vs no effect) on causality ratings.

Results Seventy-three of 4543 studies met eligibility criteria (38 RCTs and 35 non-RCTs) and 207 causality quotes were extracted. Ratings were received from 34 panelists (response rate, 89.5%). In studies in which more than 1 quote was extracted, the mean score was used. Ratings of 68 abstract quotes (35 from RCTs and 33 from non-RCTs) and 139 main text quotes (79 from RCTs and 60 from non-RCTs) were analyzed. Among the abstract quotes, the mean (SD) causality rating was 4.09 (1.56) in RCTs and 5.06 (1.13) in non-RCTs. A similar trend was found in the main text quotes (mean [SD] for RCTs, 4.18 [1.42], and for non-RCTs, 4.94 [1.21]). There was a significant main effect in the "results" variable (abstract quotes, $F = 31.42$; $P < .01$; main text quotes, $F = 51.25$; $P < .01$). No effect was found in study design and no interaction was found between the 2 independent variables.

Conclusions We failed to find statistically significant differences in the reporting of causal relationships between non-RCTs and RCTs; however, non-RCTs consistently scored higher than RCTs in the causality rating. The results suggest that quality improvement researchers may overemphasize causal inference in non-RCTs.

¹Ottawa Health Research Institute, Clinical Epidemiology Program, 1053 Carling Ave, Administration Building, Room 2-010, Ottawa, Ontario, Canada K2B 7T4, e-mail: lli@ohri.ca; ²Istituto di Igiene e Medicina Preventiva, University of Milan, Milan, Italy; ³Hospital Universitario Virgen de Valme, Seville, Spain

Web and e-Publishing

The Effect of Using e-Mail Push Technology on Readership of Articles in a General Medical Journal

George D. Lundberg and Kaytie Brown

Objective The Internet has revolutionized the distribution of information from medical journals and has irrevocably changed the reading patterns of physicians and other health care professionals. Virtually all serious primary source peer-reviewed medical journals now offer either exclusively electronic, duplicative electronic and paper, or supplemented paper and electronic forms. Receipt by the reader may be by deliberate subscription, by active

electronic linkages, via searches, or by e-mail reminders, permitted or without permission. As a part of Medscape, *MedGenMed* (<http://www.medgenmed.com>) is available free of charge to all registrants. Medscape sends newsletters to a large permission database in weekly MedPulses. In a natural, uncontrolled experiment, we tested the effect of inclusion of *MedGenMed* articles in MedPulses on the readership of those articles.

Design Forty-four articles of a variety of types published between January 1 and October 21, 2004, were studied. We tallied the number of unique users for each article. We then noted which articles had been independently chosen for MedPulse distribution by the various editors responsible for the 32 MedPulses. We compared the unique users across the number of MedPulse listings.

Results The largest number of total unique users for an article was 26 154, the smallest was 175, and the average was 5455. Five articles were not included in any MedPulses. Their average number of unique users was 1013. Forty articles were included in 1 to 29 MedPulses; their average number of unique users was 5618. The trend line was linear and positive, with wide scatter.

Conclusions Including references and hot links to *MedGenMed* articles in MedPulses had a positive effect on readership. Journals with electronic editions should consider using e-mail push permission databases to enhance readership. Additional variables will be examined in the full report.

Medscape General Medicine, 4600 Patrick Henry Dr, Santa Clara, CA 95054, USA, e-mail: glundberg@webmd.net

Prepublication Release of Journal Articles: Impact of the Electronic Medium on the Research Message

Matthew B. Stanbrook^{1,2,3} and Donald A. Redelmeier^{1,2}

Objective To determine whether prepublication release of articles via a journal's Web site influences article impact.

Design Cohort study of original research articles published in the *New England Journal of Medicine* between 1997 and 2001. The set of all such articles released in advance of print publication on the journal's Web site ("early-release articles"; n=24) was compared with 2 sets of control articles, 1 matched based on date of publication and order of listing in the journal issue ("time-matched controls"; n=93) and the other matched by disease, intervention, and study design, based on Medical Subject Headings assigned by the National Library of Medicine ("content-matched controls"; n=19). Article impact was measured by citation frequency (Web of Science, Thomson ISI) and by how frequently journal readers downloaded articles electronically (CiteTrack, HighWire Press).

Results All early-release articles were clinical studies evaluating an intervention; 62% were randomized trials and 17% were case series. Compared with both time-matched and content-matched control articles, early-release articles had approximately twice as many citations per year (47 vs 23; rate ratio, 2.03; 95% confidence interval [CI], 1.32-3.12; $P = .002$ and 50 vs 29; rate ratio, 1.71; 95% CI, 0.88-3.36; $P = .11$, respectively) and Internet downloads per year (7603 vs 3317; rate ratio, 2.29; 95% CI, 1.63-3.23; $P < .001$ and 7820 vs 3466; rate ratio, 2.26; 95% CI, 1.16-4.39;

$P = .02$, respectively). Adjustment for study quality (based on peer reviewers' ratings) and for immediacy and importance (based on a standardized evaluation by independent blinded reviewers) yielded similar estimates of the effect of early release on article citations and downloads, although differences in citations no longer reached statistical significance.

Conclusions Prepublication release appears to be associated with greater reading and citation of articles independent of study content and quality, suggesting that a journal can use electronic publication to influence how physicians perceive new research findings.

¹University of Toronto, Toronto, Ontario, Canada; ²Institute for Clinical Evaluative Sciences in Ontario, Toronto, Ontario, Canada; ³Toronto Western Hospital, Seventh Floor, East Wing, 399 Bathurst St, Toronto, Ontario, Canada M4N 3M5, e-mail: m.stanbrook@utoronto.ca

SUNDAY, SEPTEMBER 18

Authorship and Contributorship

How Intuitive Are ICMJE Criteria for Authorship? Perceptions of Deserved Authorship Among Medical Students and Physicians

Darko Hren, Dario Sambunjak, Ana Ivanis, Matko Marušić, and Ana Marušić

Objective To analyze medical students' and physicians' perceptions of research contributions as criteria for authorship in relation to the authorship criteria defined by the International Committee of Medical Journal Editors (ICMJE): (1) "conception and design of study" or "analysis and interpretation of data" or "collection and assembly of data" and (2) "drafting of the article" or "critical revision of manuscript," and (3) "final approval of the article."

Design Medical students with (n= 152) or without (n= 85) prior instruction on ICMJE criteria, graduate students and physicians attending a continuing medical education course (n= 125), and medical teachers experienced in scientific publishing (n= 112) evaluated the importance of 11 research contributions as authorship qualifications on a scale from 1 (no importance) to 4 (high importance). They also reported single contributions eligible for authorship, as well as combinations of 2 or 3 qualifying contributions. Four groups were compared on the average importance they attributed to each contribution. Hierarchical cluster analysis was performed using average importance of each contribution as well as frequency of each appearance as a single or partial authorship criterion.

Results "Conception and design of study," "analysis and interpretation of data," and "drafting of article" formed the most important cluster in all 4 groups. The effect of instruction to medical students was found for "critical revision of manuscript" and "final approval of the article." "Final approval" was a part of the least important cluster in all groups except students with instruction.

Conclusions “Conception and design,” “analysis and interpretation of data,” and “drafting of article” are ICMJE criteria for authorship recognized as most important by all participants and can be considered intuitive (ie, independent of previous instruction). “Critical revision of manuscript,” “final approval,” and “acquisition of data” are less acknowledged contributions and their significance should be taught actively to students and authors.

Croatian Medical Journal, Zagreb University School of Medicine, Salata 3b, 10000 Zagreb, Croatia, e-mail: dario.sambunjak@mef.hr

Conflict of Interest

Conflict of Interest Management at the International CPR Consensus Conference

John E. Billi,¹ Brian Eigel,² David Zideman,³ and Vinay Nadkarni,⁴ for the International Liaison Committee on Resuscitation (ILCOR) and American Heart Association (AHA)

Objective To describe a novel conflict of interest management technique implemented by the International Liaison Committee on Resuscitation and the American Heart Association (ILCOR/AHA) for the January 2005 ILCOR Consensus Conference (C2005).

Design The C2005 included 380 invited resuscitation experts. No industry representatives were invited. ILCOR/AHA received no commercial support for C2005. The ILCOR/AHA conflict of interest process applied to all participants and spanned preconference work, topic selection, systematic review, worksheet critique, worksheet Internet posting/comments, presentations, and consensus development. The conflict of interest questionnaire covered financial, business, and intellectual interests and relationships. Worksheets included mandatory disclosure of potential conflicts. Before C2005, all participants filed conflict of interest disclosure forms, reviewed for conflicts by AHA staff and ILCOR conflict of interest co-chairs, who kept written records of all concerns/actions. Corrective actions included reassigning roles to a person without significant conflicts or limiting the role to evidence review. Participants received the conflict of interest disclosure booklet listing each attendee's name, assigned conflict of interest number, institution, and basic disclosure details. A postconference survey was also conducted.

Results A total of 190 of 380 participants noted they had no conflicts to disclose. All speakers, moderators, and floor commentators stated their name/conflict of interest number. That individual's disclosures appeared immediately on a separate screen during his/her entire presentation/comment. Conflict of interest slides were updated daily. Each session's conflict of interest monitor oversaw proper disclosures and filed written reports. Moderators referred unresolved conflict of interest problems to the ILCOR conflict of interest co-chairs for rapid resolution. A conflict of interest hot-line solicited confidential concerns. A conflict of interest policy/rationale poster stimulated much discussion. Conflict of interest co-chairs investigated and recommended resolution for 8 concerns before and 12 during C2005. Once during C2005 the conflict of interest monitor terminated discussion because the commentators had commercial conflicts on issues in

the debate. The conflict of interest committee met to handle one moderator's issue. No anonymous calls were received. A post-C2005 survey (120 respondents) indicated that 90% “strongly agreed” or “agreed” that speakers' relationships with commercial entities were clearly disclosed during C2005.

Conclusions An innovative approach to disclosure and management of conflict of interest effectively guided a large, international consensus process to successful outcomes. The process was efficient, transparent, and well-received by participants.

¹University of Michigan Medical School, M7319 Med Sci I, Ann Arbor, MI 48109-0624, USA, e-mail: jbilli@umich.edu; ²American Heart Association, Dallas, TX, USA; ³Hammersmith Hospital, London, UK; ⁴Children's Hospital of Philadelphia, Philadelphia, PA, USA

Duplicate Publication

The Extent and Characteristics of Duplicate Publications in a Cohort of Meta-analyses

Veronica Yank,¹ Drummond Rennie,^{2,3} and Lisa A. Bero³

Objective Duplicate research articles are a particular problem for meta-analysts who need to avoid multiple counts of data from the same trials. One might therefore assume meta-analysts would not publish duplicate research themselves. We sought to determine the extent and characteristics of duplicate publications in a cohort of meta-analyses.

Design We reviewed meta-analyses published from January 1966 to June 2002 that evaluated antihypertensive drug treatments in nonpregnant adults. Meta-analyses were identified by electronically searching PubMed and the Cochrane Database and by hand-searching the reference lists of identified meta-analyses. Duplicate meta-analyses were defined as those that shared at least 1 author and evaluated exactly the same trials and primary outcomes measures.

Results Of 96 meta-analyses identified, 25 (26%) were duplicates, including 8 pairs and 3 sets of triplets. Given 34 opportunities for cross-referencing between duplicate meta-analyses, there were 26 (76%) cases without any cross-referencing, 3 (9%) with explicit referencing, and 5 (15%) with vague referencing. Fifteen (60%) duplicate meta-analyses did not disclose any funding source, while 3 (12%) disclosed funding by a pharmaceutical company, 6 (25%) by a government, society, or foundation, and 1 (4%) by both a pharmaceutical company and a society. Funding sources were disclosed by both meta-analyses in a pair/triplet in only 2 of 11 (18%) of related pairs/triplets. Twelve (48%) duplicate meta-analyses were published in journal supplements, 10 (40%) in regular journal issues, and 3 (12%) in books. Of the 12 journal supplements that included duplicate meta-analyses, 2 (17%) did not disclose any sponsorship and 6 (50%) were sponsored by single pharmaceutical companies, 3 (25%) by multiple pharmaceutical companies, and 1 (8%) by both pharmaceutical and governmental sources.

Conclusions A high proportion of meta-analyses on antihypertensive drug treatments are duplicates. These cross-reference one another vaguely, are inconsistent in disclosure of funding, and are often published in journal supplements.

¹University of Washington, 4411 Fourth Ave NE, Seattle, WA 98105, USA, e-mail: vyank@u.washington.edu; ²JAMA, Chicago, IL, USA; ³University of California, San Francisco, CA, USA

Ethical Concerns

Assessment of Equipoise Using a Cohort of Randomized Controlled Trials

Yuan-I Min,¹ Aynur Unalp-Arida,² Roberta Scherer,³ and Kay Dickersin⁴

Objective To evaluate whether the principle of equipoise is observed in clinical trial design.

Design An existing data set from a retrospective follow-up study (1988-1989) of 3 cohorts of initiated studies evaluating publication bias was used for this analysis. Only randomized controlled trials (RCTs) that reported a drug as the major test treatment were included (n = 111). We classified trials as favoring the test treatment if the results for the specified primary outcome were statistically significant in favor of the test treatment (based on investigators' interviews). We expect 50% of the trials to have results favoring test treatment if investigator equipoise is present with respect to selection of control type.

Results Fifty-four percent of drug trials reported findings favoring the test treatment ($P = .39$ for observed vs expected proportions, ie, 50% [observed vs expected]). Industry-sponsored trials were equally likely to report findings favoring the test treatment (59%; $P = .47$ [observed vs expected]) as nonindustry-sponsored trials (53%; $P = .53$ [observed vs expected]). Trials using a placebo or no treatment as the control were more likely to report findings favoring the test treatment (63%; $P = .03$ [observed vs expected]) compared with trials with other types of controls (32%; $P = .048$ [observed vs expected]). Industry-sponsored and nonindustry-sponsored trials were similar in the proportion of trials using a placebo or no treatment as the control (71% vs 72%); however, within this subgroup, industry-sponsored trials seemed more likely to report findings favoring test treatment (83%; $P = .02$ [observed vs expected]) than nonindustry-sponsored trials (59%; $P = .14$ [observed vs expected]).

Conclusions Our data suggest that some drug RCTs may not satisfy the equipoise principle in trial design. This is mostly observed in industry-sponsored trials that used a placebo or no treatment as the control groups. However, the number of industry-sponsored studies in our cohort is small (n = 17). Similar analysis in other cohorts of trials is necessary to generalize our results.

¹MedStar Research Institute, Department of Epidemiology and Statistics, 6495 New Hampshire Ave, Suite 201, Hyattsville, MD 20783, USA, e-mail: nancy.min@medstar.net; ²Johns Hopkins University, Center for Clinical Trials, Baltimore, MD, USA; ³University of Maryland, Department of Epidemiology and Preventive Medicine, Baltimore, MD, USA; ⁴Brown University, Department of Community Health, Providence, RI

Has Reporting on Informed Consent, Ethical Approval, Competing Interest, and Financial Support Been Improved in Randomized Controlled Trial Articles in 3 Chinese Medical Journals?

Qian Shou-chu,¹ Lv Xiao-dong,² and Liu Bin³

Objective In 1999, the rates of reporting informed consent, ethical approval, competing interest, and financial support in randomized controlled trials (RCTs) were 8.6%, 4.6%, 3.7%, and 40%, respectively, in 3 major medical journals of the Chinese Medical Association. This study is to assess whether these journals have improved their reporting after 4 years.

Design A total of 45 RCTs were published in 2004 in the *Chinese Medical Journal* (CMJ, 8 articles), the *National Medical Journal of China* (NMJC, 19), and the *Chinese Journal of Cardiology* (CJC, 18). The rates of reporting on informed consent, ethical approval, competing interest, and financial support in RCTs articles in these journals were compared with those of 1999.

Results Among these 45 articles 15 (33%) included informed consent; 14 (31%), ethical approval; 16 (36%), competing interest; and 17 (38%), financial support, which were higher than those of 1999. The CMJ did the best in reporting these elements in RCTs with the rates of 63% for informed consent, 88% for ethical approval, 50% for competing interest, and 100% for financial support.

Conclusions The 3 major medical journals of the Chinese Medical Association have improved their reporting on informed consent, ethical approval, competing interest, and financial support, but the overall improvement is not adequate.

¹Chinese Medical Journal, Chinese Medical Association, 42 Dongsixidajie, Beijing 100710, China, e-mail: qsc@ht.rol.cn.net; ²National Medical Journal of China, Beijing, China; ³Chinese Journal of Cardiology, Beijing, China

Impact Factor

Text vs Context: The Influence of the Journal on Article Impact

Matthew B. Stanbrook^{1,2,3} and Donald A. Redelmeier^{1,2}

Objective To determine whether the impact of a journal article is influenced by the journal it appears in and to estimate the magnitude of this influence.

Design In September 2001, a single article (entitled "Sponsorship, Authorship, and Accountability") appeared concurrently in 12 leading medical publications. The article's content was identical in each. We identified all subsequent citations to the article from the Institute for Scientific Information's Web of Science and classified each according to the source journal that was identified as having published the article.

Results Total citations differed by 2 orders of magnitude (range, 1-102; median, 12; interquartile range, 2-36) between the most and least frequently cited source journal. This gradient persisted over time, being observed in both the first and second year after publication. All source journals were cited by similar types of citing journals, 50% being other journals of clinical medicine. Total citations were highly correlated with each journal's 2002 journal impact

factor (Spearman correlation coefficient, 0.86, $P = .003$). Citation differences among journals were not associated with print circulation and were enhanced after exclusion of self-citations. Replication of the study using a similar, more recent simultaneous publication event in 2004 yielded virtually identical results.

Conclusions Contrary to the theory that journals act as passive conduits for scientific content, the impact of an article was strongly influenced by which journal published it. The journal impact factor appears to represent an accurate estimate of the relative ability of a journal to enhance article impact beyond the baseline contribution from article authors. These findings underscore the potential for a few high-impact journals to shape heavily the use of scientific information.

¹University of Toronto, Toronto, Ontario, Canada; ²Institute for Clinical Evaluative Sciences in Ontario, Toronto, Ontario, Canada; ³Toronto Western Hospital, Seventh Floor, East Wing, 399 Bathurst St, Toronto, Ontario, Canada M4N 3M5, e-mail: m.stanbrook@utoronto.ca

Open Access

Open Access Deposition Policies: What Does It Mean for Medical Journals?

Serena J. Cubie,¹ Gavriel A. Hollander,¹ Sarah C. Price,¹ and Richard A. Watts^{1,2}

Objective Several major funding organizations are considering mandating authors funded by them to deposit published articles in open repositories within 6 months of publication. Depending on the reactionary policy of the journal this could have an impact on subscriptions and/or submissions. The aim of this study was to assess the impact of open access deposition policies on the society-owned academic medical journal *Rheumatology*.

Design Two hundred original manuscripts were randomly selected from 537 submitted to *Rheumatology* between December 2003 and November 2004. Declared funding sources were identified from the manuscripts and classified as charity, government, university, industry, and no declared funding. Manuscripts were also categorized by publication decision.

Results Classification of the declared sources of funding for submitted and accepted manuscripts is shown in **TABLE 22**. Eighty (40%) submitted and 38 (50%) accepted manuscripts declared funding from charities, government, or universities, the main proponents of the deposition policies. Twenty-six (13%) submitted and 16 (21%) accepted manuscripts declared funding from 2 or more categories. No source of funding was declared by 111 (56%) submitted and 35 (46%) accepted manuscripts.

Conclusions Funding organization policies relating to open access deposition could have a significant impact on journals such as *Rheumatology*. It is too early to tell what effect these policies will have on journal subscriptions and/or submissions. However, it will be imperative for *Rheumatology* and other journals to closely monitor the future developments of these policies and take action accordingly.

¹*Rheumatology*, 41 Eagle St, London WC1R 4TL, UK, e-mail: scubie@rheumatology.org.uk; ²Department of Rheumatology, The Ipswich Hospital NHS Trust, Heath Road, Ipswich, Suffolk, UK

Open Source Tools for Open-Access Publishing

M. J. Suhonos and Gunther Eysenbach

Objective To review the goals of open access (to make research articles in all academic fields freely available to all) and open source software (to provide source code that is available for anyone to extend or modify). Both promote reliability and quality by supporting independent peer review and methodical distribution. Similarly, both have shown great potential to revolutionize traditional practices in the fields of academic publishing and software development by providing accessible, low-cost implementations.

Design Various open source applications (eg, OJS, ArtSys, PROS), standards (eg, NLM-DTD, XML, PDF, CC), and technologies (eg, PHP, MySQL, Java) were evaluated and appropriate ones selected to develop a prototype framework for the *Journal of Medical Internet Research* (www.jmir.org), an online, peer-reviewed, open-access journal.

Results Each was considered based on stability of vendor (or maturity of community), prospects for software support, security, flexibility of features, potential for customization, and associated costs. Those more established in the open source realm and more lightweight (eg, OJS, PHP, MySQL, CC) were found to be most effective for high-quality, low-cost, open-access publishing.

Conclusions Since the goals are similar, it seems reasonable to expect that open-access initiatives will evolve conjointly with open-source initiatives. It appears that, as appropriate tools emerge and mature, open-access publishing will benefit greatly by building on the practices and lessons learned from open-source development.

Journal of Medical Internet Research, Centre for Global eHealth Innovation, Toronto General Hospital, R. Fraser Elliott Building, Fourth Floor, Room #4S435, 190 Elizabeth St, Toronto, Ontario, Canada M5G 2C4, e-mail: geysenba@uhnres.utoronto.ca

Table 22. Classification of Funding for Manuscripts

Categories for Declared Source of Funding	No. (%) [*]	
	Submitted Manuscripts (n= 200)	Accepted Manuscripts (n= 76)
Charity	45 (23)	23 (30)
Government	48 (24)	22 (29)
University	16 (8)	9 (12)
Industry	11 (6)	6 (8)
Not declared	111 (56)	35 (46)

^{*}Percentages sum to more than 100 because some manuscripts had funding from multiple categories.

Peer Review Process

Screening Parameters for Reviewer Selection

Michael Callaham and John Tercier

Objective To identify data easily collected by journals that might predict the subsequent performance of potential candidates for peer reviewers. Almost no such predictors have been reported to date.

Design Subjects were reviewers at *Annals of Emergency Medicine* who met volume criteria and agreed to complete a questionnaire, stratified by quartiles of performance for invitation. Responses were compared with reviewer's average quality scores using a previously validated rating by editors. Variables were analyzed using univariate logistic regression with corrections for clustering.

Results A total of 116 reviewers responded with all necessary data; participation rates ranged from 77% for the lowest quartile to 93% for the highest. The 116 reviewers completed 1587 rated reviews during the 3 years, with an average rating of 3.9 (out of 5) and review time of 10 days. Questionnaire variables included years since residency training (mean, 14.5); academic rank; previous/other peer review experience (98%); member of an editorial board (45%); experience on US national grant review panel (40%); formal training in critical appraisal (outside of residency, journal club) (52%); degree in epidemiology or statistics (32%); received a grant as principal investigator (63%); and type of teaching hospital environment. None of these variables was associated with a higher average score except for years after residency (odds ratio [OR], 5.6 for < 10 years vs ≥19 years; $P < .05$) and editorial board membership (OR, 2.8; $P < .05$). Review scores declined linearly with years of experience. Results were similar regardless of how review scores were dichotomized, and also when the review (rather than the reviewer) was the unit of analysis.

Conclusions More objective selection of peer reviewers based on proven predictors might improve the process, but commonly available information about reviewer training and experience does not predict subsequent performance.

Department of Anthropology, History and Social Medicine and Division of Emergency Medicine, University of California, San Francisco, Box 0208, San Francisco, CA 94143-0208, USA, e-mail: mlc@medicine.ucsf.edu

Factors Affecting the Time From Manuscript Submission to Manuscript Acceptance

Julie Ely,¹ Mark Woolley,¹ Felicity Lynch,¹ Jane McDonald,² Leigh Findlay,¹ Yoonah Choi,¹ and Karen Woolley^{1,3}

Objective Timely publication of research is an ethical obligation. The interval between manuscript submission and acceptance affects timely publication, particularly as few manuscripts are accepted without revision. The objectives of this study were to calculate the median time between manuscript submission and acceptance for a large sample of articles from international, high-ranking, peer-reviewed medical journals and to identify factors affecting this time interval.

Design The time interval from the date of manuscript submission to acceptance was calculated for 1000 original research articles. This sample comprised 100 consecutive articles published up to January 2005 from each of 10 high-ranking (based on impact factor), international, peer-reviewed medical journals from different therapeutic areas. Analyses of variance were performed to determine the effect of various factors on time to acceptance.

Results The median time from manuscript submission to acceptance was 122 days (interquartile range, 76-195 days). The journal selected had a significant influence on this time interval ($P < .001$). A nephrology journal had the shortest time interval (75 days) and a general medicine journal had the longest time interval (210.5 days). Pharmaceutical industry sponsorship (10% of articles surveyed) was significantly associated with a longer time interval between submission and acceptance (geometric mean [days]: sponsored, 128.3; nonsponsored, 106.6; $P < .05$). No statistically significant association was detected between the time from submission to acceptance and the manuscript's primary outcome (positive vs negative finding), the type of research (clinical vs nonclinical), the author's country of origin (English as first language vs other), or the declared use of medical writing assistance.

Conclusions In this 1000-article sample, the median time between manuscript submission and acceptance was 122 days, although a wide range was evident across therapeutic areas. The journal selected and pharmaceutical industry sponsorship significantly affected the time interval from manuscript submission to acceptance.

¹ProScribe Medical Communications, 18 Shipyard Circuit, Noosaville, Queensland 4566, Australia, e-mail: kw@proscribe.com.au; ²ProScribe Medical Communications, Tokyo, Japan; ³University of Queensland, Queensland, Australia

Does Consumer Refereeing Improve the Quality of Systematic Reviews of Health Care Interventions? The Perspectives of Editors and Authors

Gill Gyte,¹ Carol Grant-Pearce,² Sonja Henderson,¹ Dell Horey,³ Sandy Oliver,⁴ and Carol Sakala⁵

Objective To determine how editors and review authors view consumer refereeing within the editorial process for preparing systematic reviews of effects of health care interventions; in particular, their assessment of the impact of consumer involvement on the quality of Cochrane reviews, and lessons for consumer involvement in health care research more generally. This information was sought to help plan a more extensive evaluation.

Design An independent researcher undertook semi-structured telephone interviews with editors, review authors, consumers, consumer coordinators, and the coordinator of a Cochrane review group. The researcher examined routine editorial documentation and undertook mapping interviews to understand aims of involving consumers in research and the Cochrane Collaboration's rationale for involving consumers as referees. A short questionnaire, asking for overall views of consumer input into the editorial process, identified review authors and consumers for telephone interview. This presentation reports results from interviews with 5 review authors selected to give diverse views, along with 4 editors and the group's coordinator. Consumer views are reported elsewhere.

Interviews were transcribed, and the main issues, impressions, and themes from each were summarized, with resulting data explored to identify themes.

Results Key points identified were that quality of consumer input was perceived to be positive; those with an overview of the review process considered that consumer input improved the final review; and earlier consumer input may be beneficial.

Conclusions This evaluation has identified key issues surrounding consumer refereeing of systematic reviews undertaken within the Cochrane Collaboration. Consumers were considered to provide important contributions, and suggestions for improvements in the process were made. Further research is planned to assess more specifically what additional contribution consumers make and whether objectively consumers improve the quality of Cochrane systematic reviews of health care interventions.

¹Cochrane Pregnancy and Childbirth Group, Liverpool Women's Hospital NHS Trust, Crown Street, Liverpool L8 7SS, UK, e-mail: ggyte@cochrane.co.uk; ²Policy Research in Engineering, Science and Technology (PREST), University of Manchester, UK; ³Australasia of the Cochrane Pregnancy and Childbirth Group and University of Newcastle, Newcastle, New South Wales, Australia; ⁴Social Science Research Unit, Institute of Education, University of London, UK; ⁵North America of the Cochrane Pregnancy and Childbirth Group and Maternity Center Association, New York, NY, USA

Quality Measurement of Reviewers' Reports by a Simple Instrument

Annemieke P. Landkroon, Hans Veecken, Peter Hart, and A. John P. M. Overbeke

Objective As again will be stated in the update of the Cochrane review on peer review, still little is known about the peer review process. One of the reasons is that peer review research is predominantly behavioral science. One of the most important tools to assess quality of reviewers' reports is an internally and externally validated scale. We tested adequacy and reliability of a simple 5-point scale that is used during years by the *American Journal of Obstetrics and Gynecology*. The objectives of this study were to validate and test a quick and simple instrument to assess review's quality and to search for a relationship between speed of returning a review by a referee and its quality.

Design The quality of 247 reviews of 119 original articles submitted to the *Dutch Journal of Medicine* was assessed using a 5-point scale (5, exceptional; 4, very good; 3, good; 2, below average; and 1, unacceptable). Every masked review was assessed independently by 3 editors of the journal. We calculated intraobserver variability by having rated 76 reviews for a second time by these editors. Interobserver variability was calculated by an intraclass correlation coefficient (ICC). We validated the 5-point scale in 2 ways: we asked the editors of 3 other medical peer reviewed journals to rate the 247 reviews using the same 5-point scale, and we sent the reviews of the original article to the authors of that article and a questionnaire consisting of 12 yes/no questions, resulting in a sum score between 0 and 12, and 1 question asking to give an overall score (between 1 and 5) for the review. In addition, the number of days between request and return of the review was noted (turn-

around time).

Results The ICC for the 3 editors was 0.62 (95% confidence interval [CI], 0.50-0.71) for the first assessment of 247 reviews. For the second assessment of 76 reviews, ICC was 0.62 (95% CI, 0.45-0.74), so there was no difference between first and second assessment. The ICC for the external editors was 0.60 (95% CI, 0.51-0.68), and the ICC for all 6 editors was 0.62 (95% CI, 0.55-0.68). The ICC for the pool of internal vs the pool of external editors was 0.86 (95% CI, 0.82-0.89). Of all 247 reviews, 240 were sent to 118 of 119 authors of the original articles, including a questionnaire. Of those 240 questionnaires, 187 (78 %) were returned. Author's response was 83% (98 of 118 authors). There was a significant correlation between sum score (mean, 7.7; median, 8.0) and overall score (mean, 3.3; median, 3.5) of the authors, Pearson correlation coefficient was 0.77 ($P < .01$). A significant correlation was found between mean total editorial quality assessment and overall score of authors: ICC was 0.28 ($P < .01$). Overall score was significantly higher for accepted than for rejected papers. Both overall and sum score were significantly higher for revision-needed than for rejected papers. Review quality as assessed by the mean of all editors was significantly higher for revision-needed than for accepted papers. No significant correlation was found between speed of return by a referee and quality of review, as calculated as the mean of the assessments of all 6 editors (Pearson correlation coefficient, -0.04 , $P = .50$). Mean turnaround time was 24 days.

Conclusions This 5-point scale has been proven to be a simple and reliable instrument for editors to assess quality of reviews. A significant correlation was found between mean editorial quality assessment (by all 6 editors) and quality determined by authors. No correlation was found between turnaround time and quality of reviews.

Nederlands Tijdschrift voor Geneeskunde (Dutch Journal of Medicine), P.O. Box 75971, 1070 AZ Amsterdam, the Netherlands, e-mail: overbeke@ntvg.nl

Behaviors of Authors and Peer Reviewers Following Change From Closed to Open System

L. Michael Posey

Objective To analyze retrospectively peer reviewers' bottom-line recommendations vis-à-vis their decisions to declare their identity during a change from a fully closed to a fully open system by a pharmacy practice journal.

Design Within pharmacy, most practice journals use fully closed systems of peer review in which identities of authors are blinded and reviewers are masked. After many years of using such a process, one journal changed to a fully open system at the beginning of 2002. However, because of mixed feelings about the change among a substantial minority of the journal's editorial advisory board, the editor gave authors and reviewers the option of requesting blinding and masking. The experiences of the first 2 years under this system were analyzed in terms of the number of authors and reviewers making this request and the bottom-line recommendations made by reviewers who agreed to be unmasked.

Results Authors of only 4 of 168 (2.4%) manuscripts that were sent to peer reviewers during 2002 and 2003 requested blinding; authors of 1 of 31 (3.2%) manuscripts that were rejected without peer review during this period requested blinding. Of the 415 critiques received from reviewers of the 168 manuscripts, 178 (43%) asked that their reviews be masked. Reviewers with more favorable recommendations tended to disclose their identity, while those recommending resubmission or rejection more frequently requested anonymity, as shown in TABLE 23 ($\chi^2_3 = 12.4$, $P < .01$).

Table 23. Blinding of Author and Peer Reviewer Identity by Peer Review Recommendation

Reviewer Recommendation	Identity Disclosed	Identity Masked
Accept	139	79
Revise and resubmit	50	54
Revise as note	32	21
Reject	16	24

Conclusions Authors adapted readily to an open peer review system, but reviewers were much more hesitant to identify themselves, particularly when recommending rejection.

Journal of the American Pharmacists Association, PO Box 6565, Athens, GA 30604-6565, USA

Continuing Medical Education Credit as an Incentive for Participation in Peer Review

Mary Beth Schaeffer, Christine Laine, and Catharine Stack

Objective Peer review for biomedical journals requires substantial effort but generally provides no tangible reward for those who volunteer to review. In 2004, peer review of articles submitted to biomedical journals became eligible for category 1 continuing medical education (CME) credit. We hypothesized that the CME program would encourage physicians to participate in peer review.

Design On November 1, 2004, the study journal began offering CME credit to reviewers who completed reviews that met minimal quality criteria within the requested time (14 days). Editors rate the quality of each review on a 5-point scale from 1 (poor) to 5 (excellent). To be eligible for CME credit, reviews must receive a rating 3 or greater. We compared the proportion of potential reviewers contacted who agreed to review in comparable calendar periods before and after the CME program. We also examined the time to review completion in each of the time periods. The final analysis will compare November 1, 2003, through April 30, 2004 (pre-CME) vs November 1, 2004, through April 30, 2005 (post-CME) and include a comparison of the quality of reviews in the pre- and post-CME periods.

Results Interim results comparing November 1 through December 31, 2004, vs the same calendar period in 2003 showed that journal staff contacted 1683 potential reviewers to obtain 368 (22%) who agreed to review in the pre-CME period compared with 1591 to

obtain 445 (28%) who agreed to review in the post-CME period ($P < .001$). The mean time to completion of review was 20.5 days in the pre-CME period and 13.5 days in the post-CME period (difference = 7 days; 95% confidence interval, 5.4-8.5; $P < .001$).

Conclusions Interim results suggest that CME credit is an effective incentive for physicians to review manuscripts for journals. A CME program that requires timely completion of reviews appears to reduce the time to completion of reviews.

Annals of Internal Medicine, 190 N Independence Mall W, Philadelphia, PA 19106, USA, e-mail: claine@acponline.org

Why Do Peer Reviewers Decline to Review? A Survey

Leanne Tite and Sara Schroter

Objective Peer reviewers are usually unpaid and their efforts not formally acknowledged. Editors of some journals experience difficulty finding appropriate reviewers who are able to complete timely reviews, resulting in publication delay. Our objective was to determine why reviewers decline to review and their opinions of reviewer incentives.

Design We conducted a Web-based survey of reviewers from 5 biomedical journals (*Archives of Disease in Childhood*, *BMJ*, *Emergency Medicine Journal*, *Gut*, and *Journal of Epidemiology & Community Health*). Questionnaire content was based on data from interviews with reviewers and feedback from the journals' online reviewing system. We randomly selected a sample of 200 reviewers (stratified by the number of times the reviewers had declined to review) from all reviewers who had been invited to review by each of the journals between January 1, 2003, and September 20, 2003.

Results We received responses from 606 of 890 (68%) active e-mail addresses. The most frequently cited factors for declining to review were conflict with other workload (197/304, 65%), having too many reviews for other journals (76/304, 25%), tight deadline for completing review (77/304, 25%), insufficient interest in the paper (53/304, 17%), and absence from work (48/304, 16%). Over half agreed that financial incentives will not be effective when time constraints are prohibitive (341/606, 56%) and that small financial incentives would not encourage reviewers to accept reviews (332/606, 55%). The most popular incentives included free access to journal content (389/606, 64%), more feedback about the quality of the review (337/606, 56%) and the outcome of the manuscript submission (347/606, 57%), appointment of reviewers to the journal's editorial board (338/606, 56%), and annual acknowledgment on the journal's Web site (342/606, 56%).

Conclusions Reviewers are more likely to accept to review a manuscript when it is relevant to their area of expertise. Lack of time is the principal factor in the decision to decline.

BMJ Editorial Office, BMA House, Tavistock Square, London WC1H 9JR, UK, e-mail: ltite@bmjgroup.com

Characteristics of Reviews for a Series of Open-Access Medical Journals: Results of a Retrospective Study

Elizabeth Wager¹, Emma C. Parkin,² and Pritpal S. Tamber²

Objective Reviewers' comments serve a number of purposes including helping authors improve their submissions and helping editors decide whether to accept them. The quality of a review can therefore be considered to comprise several distinct aspects. The objective of this study was to compare the quality of different aspects of reviews prepared for BioMed Central (a series of open-access medical journals that publish signed reviews alongside accepted articles).

Design Two hundred reviews (from 100 consecutive submissions to BioMed Central between November 2003 and April 2004) were assessed independently by 2 raters using a validated review quality instrument (RQI; Van Rooyen 1999). The RQI assesses 7 aspects of review quality using 5-point Likert scales. The raters discussed their ratings after the first 40 reviews to resolve any major discrepancies in scoring and to improve interrater reliability.

Results Scoring rank for different aspects of review quality was consistent between the 2 raters. The lowest-quality scores were associated with discussing the originality of the research (mean [SD], 1.87 [0.89]), providing evidence to substantiate comments (2.18 [0.86]), and commenting on the authors' interpretation of their results (2.28 [0.75]). Reviewers tended to perform better on providing constructive comments (2.73 [0.81]), identifying methodological strengths and weaknesses (2.41 [0.73]), and assessing the writing and organization of submissions (2.39 [0.87]). The difference in scores between the highest and lowest 3 categories was statistically significant ($P = .04$).

Conclusions Reviewers appear to perform best on aspects of reviews that focus on helping authors improve the quality of reporting but less well on aspects that help editors make selection decisions. Attempts to improve the quality of reviews should concentrate on encouraging reviewers to comment on the originality of research and on authors' interpretation of their findings and in providing evidence to support their recommendations.

¹Sideview, Princes Risborough, HP27 9DE, UK, e-mail: liz@sideview.demon.co.uk; ²BioMed Central, London, W1T 4LB, UK

Publication Bias

Constraints on Academic Freedom in Industry-Initiated Clinical Trials

Peter C. Gøtzsche,¹ Asbjørn Hróbjartsson,¹ Helle Krogh Johansen,^{1,2} Mette Haahr,¹ Douglas G. Altman,³ and An-Wen Chan⁴

Objective Constraints on the academic freedom of clinical investigators exist in industry-initiated randomized trials. We examined the prevalence and nature of such constraints.

Design Consecutive cohort study using protocols and corresponding publications for industry-initiated trials approved by the

Scientific-Ethical Committees for Copenhagen and Frederiksberg in 1994-1995, and a consecutive sample of protocols from 2004.

Results We found 44 protocols of industry-initiated trials from 1994-1995. The sponsor maintained tight control over the trial in progress in 32 cases; in 16 trials, the sponsor had access to accumulating data, and in an additional 16 trials, the sponsor could stop the trial at any time, for any reason; only in 1 case were any of these facts stated in a published trial report. It was stated in 22 of the 44 protocols that the sponsor either owned the data or needed to approve the manuscript, but such conditions for publication were not stated in any of the trial reports. An additional 18 protocols had other constraints. None of the protocols stated that investigators had final responsibility for the decision to publish data without first obtaining consent from the sponsor. We found similar constraints in a sample of 44 protocols from 2004, with increased secrecy about publication agreements between investigators and sponsors. Only 1 of the 88 protocols overall explicitly stated that there were no constraints.

Conclusions The tight sponsor control over randomized trials should be made transparent and modified to ensure research integrity, and trial protocols should be publicly available. The present state of affairs could not exist without the collaboration, or acquiescence, of academic researchers. This should be changed.

¹Nordic Cochrane Centre, Rigshospitalet, Dept 7112, Blegdamsvej 9, DK-2100, Copenhagen Ø, Denmark, e-mail: pcg@cochrane.dk; ²Institute of Medical Microbiology and Immunology, University of Copenhagen, Denmark; ³Centre for Statistics in Medicine, Oxford, UK; ⁴University Health Network, University of Toronto, Toronto, Ontario, Canada

Quality of Journal Articles

From Submission to Publication: A Study of the Tables and Figures in a Cohort of RCTs Submitted to the *BMJ*

David L. Schrager,^{1,2} Reshmi Sinha,¹ Pamela Liu,¹ Douglas G. Altman,² and Sara Schroter^{3,4}

Objective To examine the prevalence, content, and quality of tables and figures in reports of randomized controlled trials (RCTs) submitted to the *BMJ*. To compare tables and figures in the initial submission and final publication and analyze to what extent *BMJ* peer review might be responsible for any changes.

Design We obtained a cohort of RCTs submitted to the *BMJ* during May to August 2001. We conducted MEDLINE searches to determine whether rejected papers were published elsewhere. We obtained all published trials and counted and categorized the tables and figures in the initial submission and published article. Using established instruments and procedures we will analyze the quality of these tables and figures and check any *BMJ* reviewer comments to see whether changes were seemingly triggered by the review process.

Results Fifty-eight of the 75 RCTs submitted to *BMJ* have been published (12 in the *BMJ*) (TABLE 24). The number of tables and figures did not change markedly between submission and publication. Five percent of publications had no data tables, and 58% had

no data figures. Simple bar and line graphs predominated in both the original and published manuscripts (87%, 74%, respectively) compared with box and whisker plots (3%, 5%), scatterplots (8%, 5%), survival curves (0%, 11%), and receiver operating characteristic curves (3%, 5%). The number of CONSORT figures (37) was the same for submissions and publications; however, 6 papers gained and 6 papers lost their CONSORT figure.

Table 24. Characteristics of the Submitted Manuscripts and Published Articles

	RCTs Submitted to <i>BMJ</i> That Have Been Published (n= 58)				Published Articles (n= 59*)			
	No.	Mean	Median	Range	No.	Mean	Median	Range
Data tables	197	3.5	3	0-7	190	3.2	3	0-7
Data figures	39	0.67	0	0-4	38	0.64	0	0-3

*One submission produced 2 published randomized controlled trials (RCTs).

Conclusions While tables are included in most manuscripts and published articles, figures are used sparingly. The majority of figures are simple univariate plots with low data density. There appears to be little change in tables and figures from submission to publication. We are conducting analyses to confirm this and to examine their quality.

¹UCLA Emergency Medicine Center, UCLA School of Medicine, 924 Westwood Blvd, #300, Los Angeles, CA 90024-2924, USA, e-mail: schrige r@ucla.edu; ²Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, UK; ³*BMJ* Editorial Office, BMA House, London, UK; ⁴Health Services Research Unit, London School of Hygiene & Tropical Medicine, London, UK

Quality of Reporting Trials and Other Studies

Reporting Methods of Adverse Events in Randomized Controlled Trials

Curtis Sather and Jim Nuovo

Objective To describe the methods of reporting adverse events in randomized controlled trials and to assess adherence to CONSORT recommendations.

Design Five frequently cited journals were investigated: *Annals of Internal Medicine*, *BMJ*, *JAMA*, *Lancet*, and the *New England Journal of Medicine*. For each journal, all randomized controlled trials conducted on the use of a medication were selected from January 2000 through June 2003. All issues of each journal were reviewed manually. Information retrieved included any mention of adverse events in the abstract, methods, results, or discussion section of the article or inclusion of adverse events in tables or figures. We also catalogued whether there was a separate subheading in the results section for reporting adverse events. For adherence to CONSORT recommendations, each article was assessed for estimates of the frequency of the main severe adverse events and reasons for treatment discontinuation and an operational definition for measures of the severity of adverse events. Subsequent reports containing the same data set from a prior study were excluded from analysis.

Results There were 521 eligible articles. Explicit reporting of adverse events was found in 63% of abstracts (range, 47%-66%), 73% of methods (range, 51%-81%), 89% of results (range, 80%-95%), 21% of figures (range, 17%-52%), and 48% of tables (range, 31%-49%). There was a separate subheading for adverse events in 46% (range, 22%-64%) of the eligible articles. Adherence to noted CONSORT recommendations was present in 62% of eligible articles.

Conclusions There is variation among authors and journals as to the location of reporting adverse events and the means by which it is done. Adherence to current CONSORT recommendations for reporting adverse events is suboptimal. Efforts should be undertaken to help improve the method and form of reporting adverse events.

Department of Family & Community Medicine, UC Davis, 4860 Y St, Suite 2300, Sacramento, CA 95817, USA, e-mail: james.nuovo@ucdmc.ucdavis.edu

Grading the Evidence of Published Papers for the Benefit of Clinicians

James R. Scott,^{1,2} Rebecca Rinehart,¹ and Catherine Y. Spong¹

Objective Various classifications have been proposed to rate the quality and strength of evidence for published studies to guide physicians and benefit patient care. There are currently over 100 grading systems, but most are complicated, cumbersome, and impractical. Currently, few journals provide the level of evidence of published papers for their readers. The purpose of this study was to develop a level of evidence grading system that is useful for clinicians.

Design During a 6-month pilot study, editorial board members of *Obstetrics & Gynecology* assigned the level of evidence for articles using the American College of Obstetricians & Gynecologists (ACOG) Practice Bulletins classification system (n=47): I, properly designed randomized controlled trial; II-1, well-designed controlled trial without randomization; II-2, well-designed cohort or case-control study; II-3, multiple time series or dramatic results; III, expert opinion and other descriptive studies. When the manuscript was sent for revision, the author was asked to verify or clarify the rating, and if accepted, the editor evaluated the rating and assigned the final grade.

Results In the pilot study, there was 80% agreement between the editorial board reviewer and author and 79% agreement between the editorial board reviewer and editor. However, certain types of papers routinely submitted did not fit easily into the classification system. Most discrepancies involved 2 types of studies: well-designed controlled trial without randomization (II-1) and large case series with no control group (II-3 vs III). Modifications were made in those categories by defining II-1 as a randomized controlled trial not blinded and by placing large case series with no control group under II-3. Publishing the level of evidence was also limited to original research studies. Since January 2004, we have published the level of evidence for every original research article at the end of the abstract. A follow-up survey indicated that our readers understand and value the published grading system.

Conclusions This has proven to be a convenient and well-received grading system for reviewers, editors, and readers. We suggest that editors of medical journals agree on a uniform system to provide a level of evidence for all clinical research studies published.

¹*Obstetrics & Gynecology*, Washington, DC, USA; ²Department of Obstetrics and Gynecology, University of Utah School of Medicine, 423 Wakara Way, Suite 201, Salt Lake City, UT 84108-1242, USA, e-mail: jscott@hsc.utah.edu

Statistics

Independent Re-analyses of Identical Data Sets: Implications for Peer Review

Penelope J. Greene

Objective Clinical studies are typically analyzed by 1 statistician or a group of statisticians who have discretion selecting among various statistical procedures and options. The research objective was to compare conclusions from independent reanalyses of identical data sets.

Design Twelve statisticians independently reanalyzed original data sets from 6 published studies, without knowing published outcomes. The statisticians knew this research compared independent statistical conclusions. All data sets involved 2-treatment comparisons. Two 3×3 Greco-Latin square blocks were used, with 6 statisticians each reanalyzing 3 data sets in a block, with the ordering of the 3 data sets varied. For each of the 2-treatment comparison reports, each statistician included a statement including which outcome was statistically significantly “better” than the other, or that there was no statistically significant difference between the 2.

Results Of 14 original key 2-treatment comparison questions (from the published studies), there were only 4 comparisons for which all 6 statisticians reached the same conclusion (the same outcome significantly better, or no significant difference). For the other 10 questions, there was some disagreement. This included 3 questions for which at least 2 of the 6 statisticians concluded that different outcomes were statistically significantly “better” than the other. (For example, at least 1 statistician concluded that treatment “A” was significantly more effective than treatment “B,” and at least 1 statistician concluded that “B” was significantly more effective than “A.”) The discordant conclusions reflected the use of different statistical techniques, none of which would reasonably be considered “incorrect.”

Conclusions The professional analytic discretion exercised by statisticians can affect the research outcomes in even relatively “simple” 2-treatment comparison experiments, when the exact data are reanalyzed independently. This possibility is unlikely to be detected during any reasonable peer review process and has important implications for interpretations of research results.

Department of Nutrition, Harvard School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA, e-mail: penelope_greene@harvard.edu

Reliability of 3 Types of Research Methodologies in Obstetrics

Kerry M. McMahon, D. Yvette LaCoursiere, and James R. Scott

Objective A hierarchy exists for evidence based on the research methodology used in the study. Randomized controlled trials (RCTs) receive the most credence among all research designs. Survival analysis is a tool that allows differentiation of both proportional and temporal differences in outcomes. Thus using survival analysis, we set out to test the null hypothesis that there is no significant difference in research outcomes and length of truth survival between 3 types of methodologies: RCT, nonrandomized prospective study, and meta-analysis.

Design We identified 2 topics in the obstetric literature of sufficient longevity and quantity among the 3 types of methodologies: preterm labor and preeclampsia. A PubMed search was limited to English-language and peer-reviewed journals, which included predominantly *JAMA*, *Lancet*, *American Journal of Obstetrics and Gynecology*, and *Obstetrics & Gynecology*. From 1980-2005, we randomly selected 50 RCTs, 50 observational studies, and 50 meta-analyses for each topic. The abstracts were used to identify the authors’ research conclusion regarding management. Forty-two board-certified perinatologists will review a random selection of 50 conclusions and judge the statements to be true, false, or obsolete. The gold standard for truth will be defined by expert opinion of current medical knowledge.

Results Complete results are pending. For those conclusions found to be false or obsolete, a review of the literature will be performed to identify the year that the conclusion was refuted. The data will be stratified by methodologic type and survival curves for “truth” will be generated. Survival analysis permits a temporal evaluation between the 3 types of methodologies.

Conclusions Studies on the reliability of evidence-based medicine techniques and evaluation of research methodology in obstetrics are limited. This will be the first study to attempt to quantify the validity and duration of “truth” of research conclusions in obstetrics based on the type of methodology used.

Department of Obstetrics and Gynecology, University of Utah School of Medicine, 30 North 1900 E, Department of OB/GYN 2B200, Salt Lake City, UT 84132, USA, e-mail: kerry.mcmahon@hsc.utah.edu

Ratio Measures in Leading Medical Journals: Where Are the Underlying Absolute Risks?

Lisa M. Schwartz,^{1,2} Steven Woloshin,^{1,2} Evan L. Dvorin,² and H. Gilbert Welch^{1,2}

Objective Although ratio measures (eg, relative risk = 2) are a standard way to compare outcomes in 2 groups, without the underlying absolute risks (eg, 1-year risk of death was 0.002% vs 0.001%), they may exaggerate the magnitude of the effect. We examined the accessibility of absolute risks for ratio measures presented in leading medical journals.

Design We searched MEDLINE for abstracts with ratio measures (eg, odds ratio, relative risk, risk ratio, rate ratio, hazard ratio) in articles published by the *Annals of Internal Medicine*, *BMJ*, *JAMA*,

Journal of the National Cancer Institute, Lancet, and New England Journal of Medicine between June 2003 and May 2004. We limited our search to the 228 articles with designs in which absolute risks are directly calculable (64 randomized trials, 164 cohort studies). Each author coded a subset of articles; 30 were double coded to establish interrater reliability ($\kappa = 0.7-1.0$).

Results The average abstract had 3 ratio measures (range, 1-14). In an analysis restricted to the first ratio measure that appeared in each abstract, 70% (ranging from 51%-79% across the journals) did not include the underlying absolute risks in the abstract. Among these, 47% provided the absolute risks elsewhere in the article (in the text, a table, or a figure); 34% did not provide absolute risks but these could be calculated based on data presented, and in 19%, absolute risks were not provided and could not be calculated. Observational studies were less likely than randomized trials to provide the absolute risks in the abstract (19% in observational studies vs 58% in randomized trials; relative risk = 0.33; 95% confidence interval, 0.22-0.48).

Conclusions Articles using ratio measures often fail to make the underlying absolute risks easily accessible to readers, particularly in the case of observational studies. To improve their accessibility, absolute risks should routinely be included in the abstract adjacent to the corresponding ratio measure.

¹VA Outcomes Group, (111B), Department of Veterans Affairs Medical Center, White River Junction, VT 05009, USA, e-mail: lisa.schwartz@dartmouth.edu; ²Dartmouth Medical School, Hanover, NH, USA

Web and e-Publishing

Inviting Conversation: Engaging Diverse Individuals and Groups in an Interactive Online Forum on Published Research

Laura A. McLellan,¹ Robin S. Gotler,¹ and Kurt C. Stange^{1,2}

Objective To evaluate use of invited comments to stimulate online discussion of published research by readers, authors, and those potentially affected by the findings.

Design We evaluated the process and outcome of a prompted online discussion during the 2 years since the launch of a new primary care research journal in May 2003. Prior to publication of manuscripts, editors and authors identified individuals and constituencies potentially affected by the research. Those identified received an embargoed PDF manuscript and a request to comment in an article-specific discussion group on the journal's full-text free Web site. Authors and readers were encouraged to participate as well. The editors summarized the online discussion in a regular editorial feature. We tabulated the number of comments per article and stratified by whether the discussant had been invited to participate. We conducted a content analysis of the discussion to summarize types of comments.

Results A total of 169 articles and editorials generated 583 comments (mean, 3.45 comments per article; range, 0-63), accounting for 3.3% of journal Web site hits. Of these, 291 comments (50%) were from invited commentators and 64 (11%) were from authors. Discussants included clinicians, researchers, patients, advocacy groups, policy makers, and educators. Some articles with salient

content for the general readership generated particularly robust discussions, as did articles that activated specific groups. The content of the discussion has been remarkably thoughtful, with many comments approaching the sophistication of editorials. Four categories of comments emerged from our content analysis: interpretation that criticizes or contextualizes the article based on the reader's experience or knowledge, exhortation for advocacy or action, questions for the authors or further research, and use of the article as a catalyst for discussion of a related subject.

Conclusions Prompting comments in an online discussion of published research articles can stimulate thoughtful discourse that engages diverse participants.

¹*Annals of Family Medicine*, Cleveland, OH, USA; ²Case Comprehensive Cancer Center, Case Western Reserve University, 10900 Euclid Ave, Cleveland, OH 44106, USA, e-mail: annfammed@case.edu

The Use of the World Wide Web by Medical Journals

David L. Schrager,^{1,2} Sripha Ouk,¹ and Douglas G. Altman²

Objective The 2-page to 7-page print journal article has been the standard for 200 years, yet this format severely limits the amount of detailed information that can be conveyed. The World Wide Web provides a low-cost option for posting extended text and supplementary information. It also can enhance the experience of journal editors, reviewers, readers, and authors through added functionality (eg, online submission and peer review, postpublication critique, e-mail notification of table of contents). Our aim was to characterize ways that journals are using the Web in 2005 and note changes since 2002.

Design We iteratively developed a taxonomy of 55 ways that the Web might be used by medical journals. For items related to the Web publication of supplementary materials we will compare print and electronic issues of randomly selected 2002 and 2005 issues of journals selected on the basis of their scientific impact factor. For items related to use of the Web to enhance functionality we will review print journals and journal Web sites, and we will interview journal personnel in early 2005.

Results To date we have examined March 2002 issues of 5 general medicine and 6 specialty journals. Of 322 articles, 13 (4%) (all in *Pediatrics*) were published only on the Web. Fourteen articles in *BMJ* had longer versions on the Web than in print and 8 articles had online-only supplementary material (*BMJ*, *New England Journal of Medicine*). Four journals (86 articles [27%]) allowed online responses. There were 165 of these in 38 articles (44%) and authors responded in 9 (24%).

Conclusions Few articles in our initial 2002 sample had Web-only supplementary material. We expect such material to be more prevalent in the 2005 sample. We will characterize the different ways that journals are using the Web, thereby developing a compendium of current uses and assessing how common each usage is.

¹UCLA Emergency Medicine Center, UCLA School of Medicine, 924 Westwood Blvd, #300, Los Angeles, CA 90024-2924, USA, e-mail: schrager@ucla.edu; ²Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, UK

Sponsors

ELSEVIER

www.elsevier.com

Elsevier is a world-leading, multiple-media publisher of scientific, technical, and health information products and services, with 7,000 employees in 99 locations around the globe. Publisher of more than 20,000 products and services, including journals, books, electronic products, services, databases, and portals serving the global scientific, technical, and medical (STM) communities, its mission is to fuel a continuous cycle of exploration and discovery and inspire meaningful action. Located in North America, Europe, Middle East, Asia, Australia, Africa, and Latin America, Elsevier is part of the Reed Elsevier Group plc, a leading international publisher and information provider.

WOLTERS KLUWER HEALTH

www.lww.com

Wolters Kluwer Health is a leading provider of information for professionals and students in medicine, nursing, allied health, pharmacy, and the pharmaceutical industry under the major brands of Lippincott Williams & Wilkins, Facts & Comparisons, Ovid Technologies, Medi-Span, SKOLAR, and the pharmaceutical information provider Adis International. Medical and nursing journals published by the company are published under the Lippincott Williams & Wilkins imprint and are distributed electronically via the Ovid Technologies platform. Outstanding titles include the periodicals program of the American Heart Association and of the American Academy of Neurology.

AMERICAN HEART ASSOCIATION

www.americanheart.org

Since 1924 the American Heart Association has helped protect people of all ages and ethnicities from the ravages of heart disease and stroke. These diseases, the nation's number 1 and number 3 killers, and other cardiovascular diseases, claim nearly a million American lives a year. The association invested more than \$439 million in fiscal year 2003-2004 for research, professional and public education, and advocacy so people across America learn what they can do to reduce their risk and live stronger, longer lives. The American Heart Association's 5 scientific journals are recognized as the world's premier cardiovascular and stroke publications.

AMERICAN MEDICAL ASSOCIATION

www.ama-assn.org

The American Medical Association (AMA) helps doctors help patients by uniting physicians nationwide to work on the most important professional and public health issues. Established in 1847, the AMA is the largest member-based organization for physicians in the United States. The AMA is governed by a Board of Trustees, which ensures that the AMA remains focused on its core purpose—to promote the art and science of medicine and the betterment of public health—and a House of Delegates, the organization's principal policy-making body.

NEW ENGLAND JOURNAL OF MEDICINE

www.nejm.org

The *New England Journal of Medicine* is the oldest continuously published medical journal in the world, with over 500,000 readers in more than 177 countries. Founded in Boston in 1812, the *New England Journal of Medicine* publishes the very best information from research, at the interface of biomedical science and clinical practice, in an understandable and clinically useful format. The *New England Journal of Medicine* is owned and published by the Massachusetts Medical Society.

ANNALS OF INTERNAL MEDICINE

www.annals.org

Established in 1927 by the American College of Physicians (ACP), the *Annals of Internal Medicine* is the leading journal for studies in internal medicine. The purpose of the journal—to promote excellence in the clinical practice of internal medicine—is supported by presentation of a wide variety of experimental and clinical subject matter in the Article, Brief Communication, Update, and Review formats. And to support the belief that physicians should also be well-informed citizens of both the medical community and society at large, *Annals* offers background and discussion of issues that influence both physicians and patients.

THE COMMONWEALTH FUND

www.cmf.org

The Commonwealth Fund is a private, independent foundation dedicated to promoting a high-performing health care system that improves access, quality, and efficiency, particularly for society's most vulnerable populations. The Commonwealth Fund supports research, makes grants to improve health care practice and policy, and has an international program designed to stimulate innovative health care policies and practice in the United States and other industrialized countries.

THE JOSIAH MACY, JR FOUNDATION

www.josiahmacyfoundation.org

The Josiah Macy, Jr Foundation is a privately endowed philanthropy located in the borough of Manhattan, New York City. The Foundation supports programs designed to improve the education of health professionals in the interest of the health of the public and to enhance the representation of minorities in the health profession.

THE ROCKEFELLER UNIVERSITY PRESS

www.rockefeller.edu/rupress

The Rockefeller University entered publishing in 1905 with the production of *The Journal of Experimental Medicine*, coedited by the institution's first head, Simon Flexner. The Publications Department at Rockefeller was established in 1910 and expanded its program to include monograph series and "Studies from the Rockefeller Institute for Medical Research." It published *The Journal of Biological Chemistry* from 1914 to 1925 and founded 2 journals of its own: *The Journal of General Physiology* in 1917 and *The Journal of Biophysical and Biochemical Cytology* (later changed to *The Journal of Cell Biology*) in 1955. The Department published *The Journal of Clinical Investigation* for the American Society for Clinical Investigation from 1967-1998.

UNICEF / UNDP / WORLD BANK / WHO

SPECIAL PROGRAMME FOR RESEARCH AND TRAINING IN TROPICAL DISEASES

www.who.int/tdr

The Special Programme for Research and Training in Tropical Diseases (TDR) is an independent global programme of scientific collaboration. Established in 1975 and co-sponsored by the United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP), the World Bank, and the World Health Organization (WHO), the TDR aims to help coordinate, support, and influence global efforts to combat a portfolio of diseases of the poor and disadvantaged. The WHO is the Executing Agency of TDR, which is based at WHO headquarters in Geneva, Switzerland.

WORLD ASSOCIATION OF MEDICAL EDITORS (WAME)

www.wame.org

The World Association of Medical Editors (WAME) is a voluntary association of editors from many countries who seek to foster international cooperation among editors of peer-reviewed medical journals. With more than 1200 members representing 764 journals from 85 countries, WAME helps medical journal editors fulfill their responsibility to ensure that reports of medical research provide valid information in a form that is readily accessible by researchers, medical practitioners, and others by providing educational resources and a forum for discussion of issues in research publication. In particular, WAME aims to assist editors in developing countries and editors of small journals, who often face difficulties obtaining high-quality manuscripts, may lack formal training in editing, and have limited finances and access to publication expertise.

NEDERLANDS TIJDSCHRIFT VOOR GENEESKUNDE (DUTCH JOURNAL OF MEDICINE)

www.ntvg.nl

The *Nederlands Tijdschrift voor Geneeskunde* (Dutch Journal of Medicine) is the only general medical scientific weekly in the Netherlands. It was founded in 1856 and published the first issue in January 1857. The journal has approximately 30,000 subscribers, reaching 75% of Dutch doctors. It is published in Dutch and is therefore a national journal, with quite a lot of impact.

Exhibitors

Allen Press, Inc.

810 East 10th St
Box 368
Lawrence, KS 66044
Anna Jester
785-843-1234
ajester@allenpress.com

Aries Systems Corporation

200 Sutton St
North Andover, MA 01845
Sandy Holmes
978-975-7570
sholmes@ariessys.com

Burg Translations Inc.

29 South LaSalle St
Suite 936
Chicago, IL 60603
Bill Branda
312-263-3372
marketing@burgtranslation.com

Cadmus Communications

1801 Bayberry Ct
Suite 200
Richmond, VA 23226
Mark Anderson
804-287-3664
AndersonM@cadmus.com

Chalex Corp

53 Greenwood Shoales
Grasonville, MD 21638
Murray Oles
410-827-6070
gmurray@chalexcorp.com

Council of Science Editors (CSE)

CSE Headquarters
c/o Drohan Management Group
12100 Sunset Hills Rd
Suite 130
Reston, VA 20190
Kathy Hoskins
703-437-4377
Fax 703-435-4390
CSE@CouncilScienceEditors.org

eJournalPress

4641 Montgomery Ave
Suite 515
Bethesda, MD 20814
Brynn Gipe
301-961-6033
support@ejournalpress.com

Inera Inc.

815 Washington St
Suite 3
Newton, MA 02460
Elizabeth Blake
617-969-5287
lblake@inera.com

Mark Logic Corporation

2000 Alameda de las Pulgas
Suite 100
San Mateo, CA 94403
Traci White
650-655-2300
traci.white@marklogic.com

NRC Research Press-National

Research Council Canada
1200 Montreal Rd
Ottawa, ON K1A 0R6 Canada
Jamie Gregoire
613-993-9085
jamie.gregoire@nrc.gc.ca

ScholarOne, Incorporated

375 Greenbrier Dr
Suite 200
Charlottesville, VA 22901
Sue Cifelli
434-817-2040 ext 131
sue.cifelli@scholarone.com

JAMA (JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION)

www.jama.com

JAMA is a weekly, international, peer-reviewed, general medical journal with a print circulation of 360,000 in more than 150 countries, a much wider online reach, and 12 international editions in 11 languages. *JAMA*'s mission is to promote the science and art of medicine and the betterment of the public health. The editors and publishers of *JAMA* also publish 9 *Archives* Journals. *JAMA*, which is editorially independent of the American Medical Association, was founded in 1883 and has been published consecutively for 122 years.

The organizers wish to thank the following individuals who contributed significantly to the planning and support of this Congress: Jeni Reiling, Rosa Miranda, Sharon Kremkau, Kim Westman, Marla Hall, Maggie Mills, Delia Rico, Cassio Lynn, Alison Burke, Ronna Henry Siegel, Vivienne Williams, Kay Carow, Peter Kim, Michelle Kurzynski, Terri Hanley, Stacy Christiansen, Kim Campbell, Joy Jaeger, Phil Sefton, Heather Shebel, Beverly Stewart, Carole Felty, Jann Ingmire, and Jim Michalski.

BMJ PUBLISHING GROUP

www.bmjpg.com

The *BMJ* Publishing Group produces a wide range of journals, Web sites, and support for doctors, including the weekly *BMJ*, over 20 specialty journals, *Clinical Evidence*, and several local editions of the *BMJ* that are published overseas. The *BMJ* Publishing Group is owned by the British Medical Association (BMA), the professional association for doctors in the UK. The *BMJ* was founded in 1840 and has been published continuously ever since. Today, it is one of the world's leading international medical journals.

Notes

With Sponsorship From

Elsevier

www.elsevier.com

Wolters Kluwer Health

www.lww.com

American Heart Association

www.americanheart.org

American Medical Association

www.ama-assn.org

New England Journal of Medicine

www.nejm.org

Annals of Internal Medicine

www.annals.org

The Commonwealth Fund

www.cmf.org

The Josiah Macy, Jr Foundation

www.josiahmacyfoundation.org

The Rockefeller University Press

www.rockefeller.edu/ruPress

UNICEF/UNDP/World Bank/WHO

**Special Programme for Research
and Training in Tropical Diseases**

www.who.int/tdr

World Association of Medical Editors (WAME)

www.wame.org

***Nederlands Tijdschrift voor Geneeskunde*
*(Dutch Journal of Medicine)***

www.ntvg.nl

JAMA
&
ARCHIVES
JOURNALS

BMJ
Publishing
Group